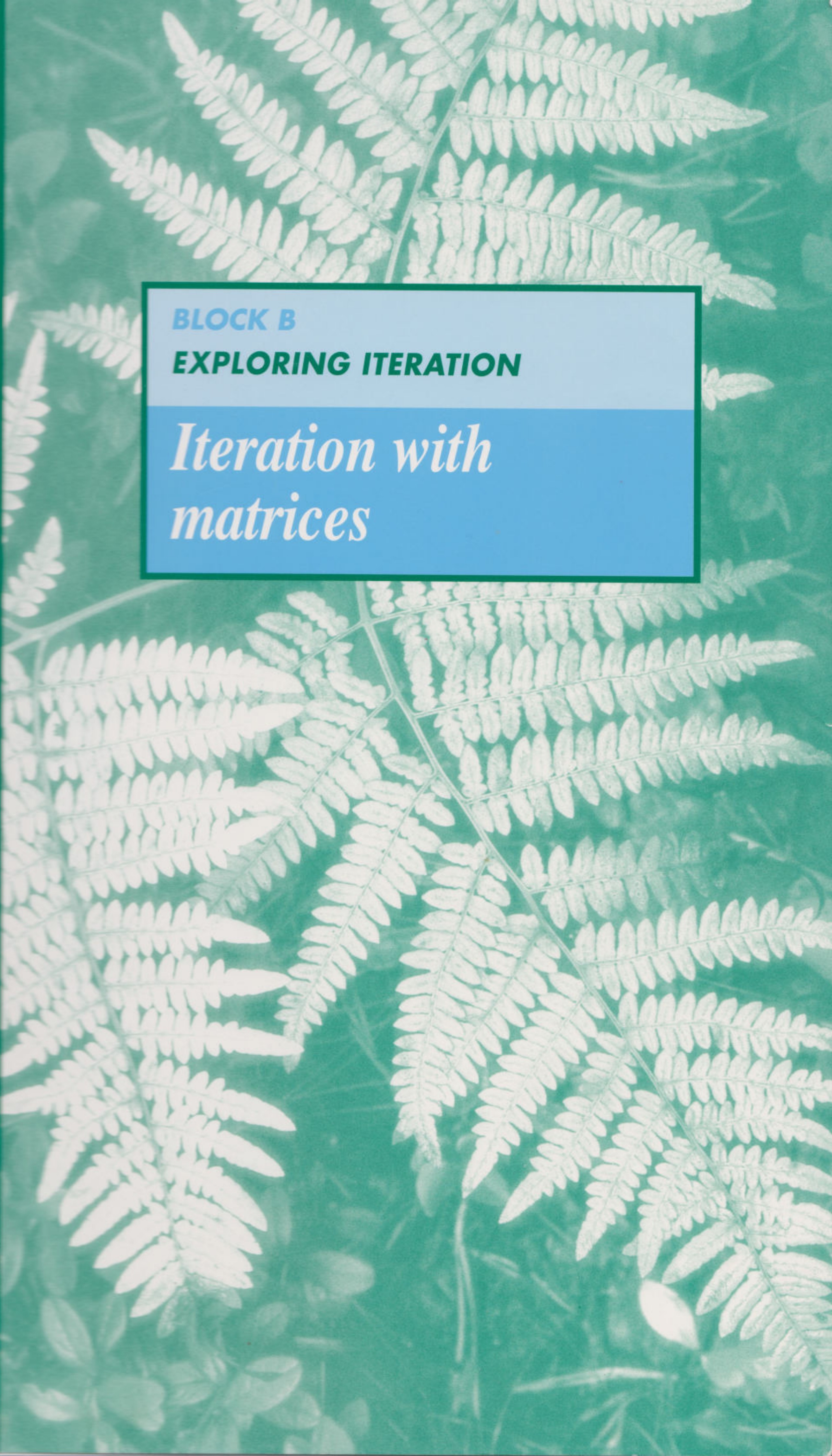
MS221 Chapter B3

The Open University

A second level interdisciplinary course

Exploring Mathematics



BLOCK B

EXPLORING ITERATION

Iteration with matrices

CHAPTER B3

The Open University

A second level interdisciplinary course

Exploring Mathematics

CHAPTER

B3

BLOCK B

EXPLORING ITERATION

Iteration with matrices

Prepared by the course team

About this course

This course, MS221 *Exploring Mathematics*, and the courses MU120 *Open Mathematics* and MST121 *Using Mathematics* provide a flexible means of entry to university-level mathematics. Further details may be obtained from the address below.

MS221 uses the software program Mathcad (MathSoft, Inc.) to investigate mathematical concepts and as a tool in problem solving. This software is provided as part of the course.

This publication forms part of an Open University course. Details of this and other Open University courses can be obtained from the Student Registration and Enquiry Service, The Open University, PO Box 197, Milton Keynes, MK7 6BJ, United Kingdom: tel. +44 (0)870 333 4340, e-mail general-enquiries@open.ac.uk

Alternatively, you may visit the Open University website at http://www.open.ac.uk where you can learn more about the wide range of courses and packs offered at all levels by The Open University.

To purchase a selection of Open University course materials, visit the webshop at www.ouw.co.uk, or contact Open University Worldwide, Michael Young Building, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom, for a brochure: tel. +44 (0)1908 858785, fax +44 (0)1908 858787, e-mail ouwenq@open.ac.uk

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 1997. Second edition 2002. Reprinted 2003. Reprinted with corrections 2005.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed and bound in the United Kingdom by The Charlesworth Group, Wakefield.

ISBN 0 7492 4035 1

2.3

Contents

Study guide

There are five sections to this chapter. They are intended to be studied consecutively in five study sessions.

Section 1 requires the use of an audio cassette player, and Section 5 requires the use of the computer together with Computer Book B.

The pattern of study for each session might be as follows.

Study session 1: Section 1.

Study session 2: Section 2.

Study session 3: Section 3.

Study session 4: Section 4.

Study session 5: Section 5.

Each session will require two to three hours to study, the longest being the second.

In addition, there is an *optional* video band associated with this chapter, in which an illustration is given of the possible uses of certain linear transformations in design. The best time to watch this band is after you have studied Section 4.

Before studying this chapter, you should be familiar with the following topics:

- ◇ the concept of a fixed point of a function;
- ◇ linear transformations, in particular, rotations, reflections, scalings and shears;
- ◇ the matrix representation of a linear transformation;
- ◇ the determinant test for the invertibility of a 2×2 matrix;
- ◇ matrix algebra including matrix multiplication and the formula for the inverse of a 2×2 matrix, involving its determinant;
- ◇ the long-term behaviour of the geometric sequence r^n, and the evaluation of limits of ratios involving geometric sequences.

> The optional Video Band B(iii) *Algebra workout – Eigenvectors* could be viewed at any stage during your study of this chapter.

Introduction

In this chapter two concepts, iteration of functions from Chapter B1 and linear transformations from Chapter B2, are drawn together into the study of the iteration of linear transformations.

The main theme is the further study of linear transformations, but affine transformations, which were introduced in Chapter B2, are also investigated briefly.

Any function f with rule $\mathbf{x} \longmapsto \mathbf{Ax}$, where $\mathbf{A}$ is a 2×2 matrix and $\mathbf{x}$ is a vector, represents a linear transformation. You will see that a geometric understanding of such functions provides a key to understanding the long-term behaviour of their iterations. In particular, it turns out to be important to identify certain sets of points that remain unchanged when a linear transformation is applied.

This chapter begins by studying the geometry of certain types of linear transformations, and looking at the points which remain fixed under these transformations. We also look at lines which remain unchanged under these transformations; on such lines, the individual points might not be fixed by the transformation, but the image of the entire line is the same as the original line. We call such a line an *invariant line*. Invariant lines through the origin are of particular interest. In Section 1, we look at these invariant lines from a geometric perspective for certain types of linear transformation, but we soon find that we need to investigate them algebraically. Section 2 takes an algebraic approach which allows us to find such invariant lines for many linear transformations represented by 2×2 matrices.

In Section 3, you will see how to write certain 2×2 matrices $\mathbf{A}$ in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal 2×2 matrix. This process, called *diagonalisation*, gives a neat and efficient way of finding powers of a matrix.

In Section 4, we use the results from Sections 2 and 3 to explore the iteration of linear transformations. It turns out that the long-term behaviour of sequences of points generated by iteration of linear transformations can often be described by using properties of the corresponding matrix.

In Section 5, the computer is used to investigate patterns in the iteration of linear and affine transformations. A description is given of how iteration can lead to pictures of natural objects such as ferns.

1 Fixed points and invariant lines

To study this section you will need an audio cassette player and Audio Tape 2.

When a typical linear transformation is applied to the plane, most points are moved to a new position. As a consequence, the geometrical properties of a figure (shape, angles, distances, etc.) are altered when the linear transformation is applied. However, certain things are not altered by the application of a linear transformation – for example, the origin is unchanged, and some straight lines may also remain unchanged. In this section we investigate points and lines which are unchanged under certain basic linear transformations.

Recall that a linear transformation maps the origin to itself and maps straight lines to straight lines.

In the audio tape, we use our geometric understanding of rotations, reflections, scalings and shears to determine the points and lines which remain unchanged when these types of linear transformation are applied.

A point which remains unchanged under a linear transformation is called a *fixed point* of that transformation. The first half of the audio tape considers fixed points of particular linear transformations. The second half of the audio tape considers lines which remain unchanged under the same linear transformations. These lines are called *invariant lines*.

Recall from Chapter B1, Section 1, that a fixed point of a real function f is a real number a such that $f(a) = a$.

Now listen to Audio Tape 1, Band 4 'Fixed points and invariant lines'.

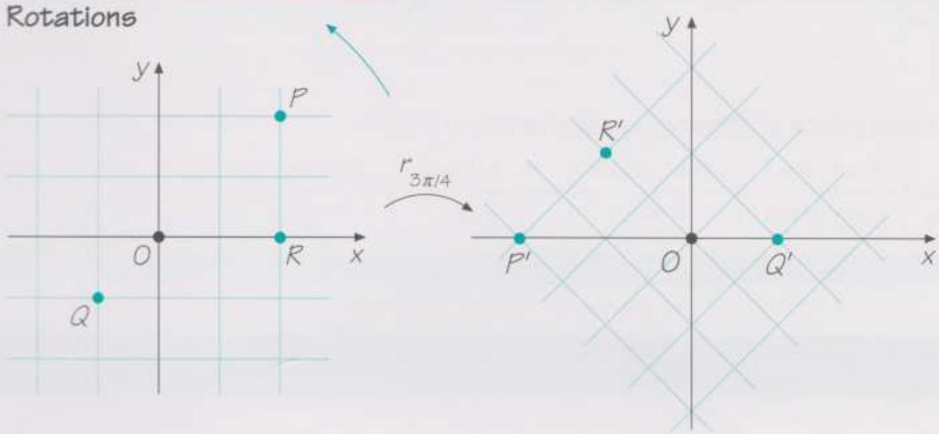
Frame 1

Fixed Points

A **fixed point** of a linear transformation is a point which is equal to its image under the linear transformation.

Frame 2

Rotations



All points move except the origin.

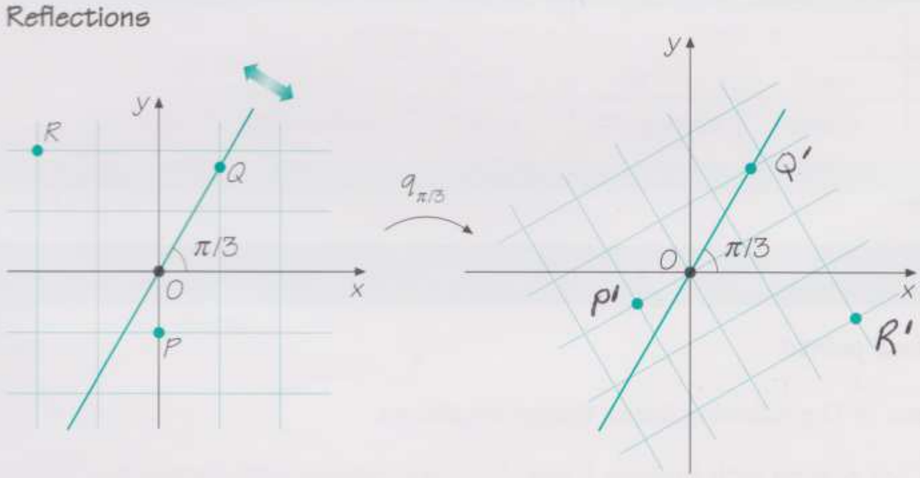
Fixed points of the rotation r_θ:

Angle	$\theta \neq 2k\pi$	$\theta = 2k\pi$
Fixed points	O	every point

$k \in \mathbb{Z}$

Frame 3

Reflections



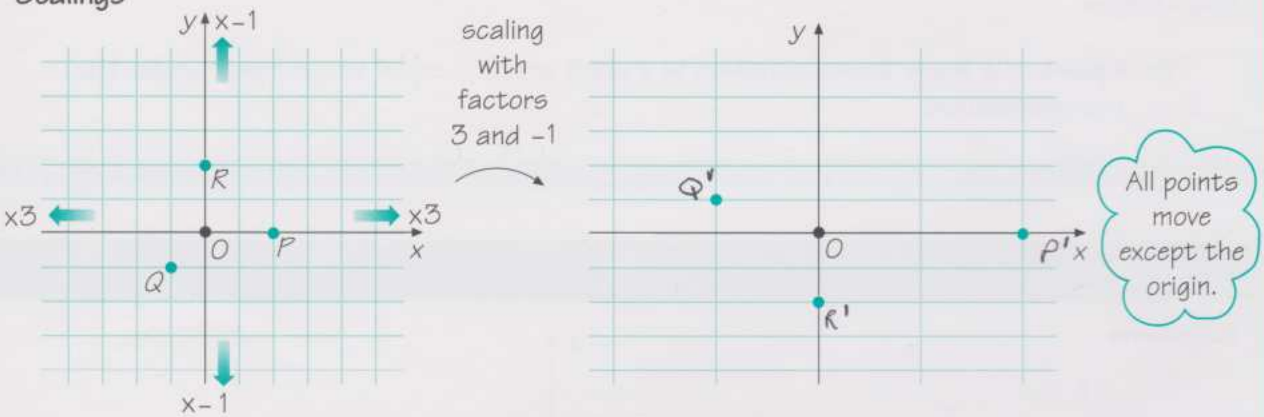
All points on the axis of reflection are fixed.

Fixed points of the reflection q_θ: Q

All points on the axis of reflection

Frame 4

Scalings

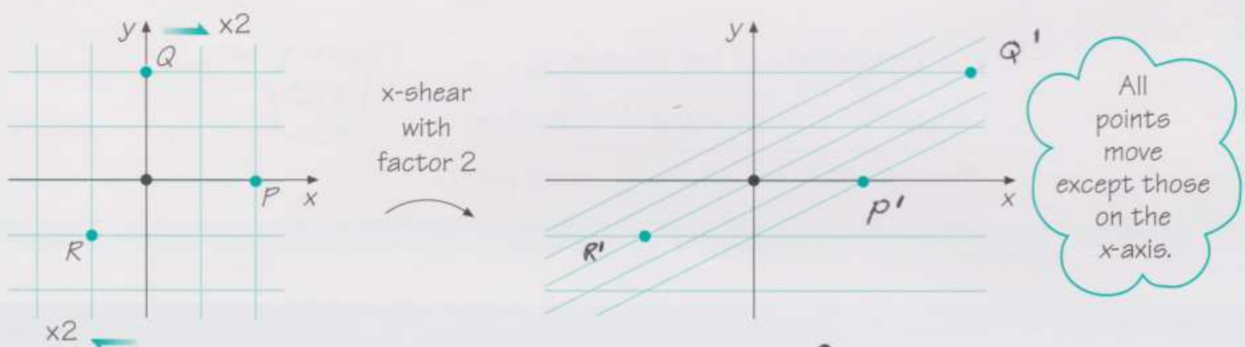


Fixed points of the scaling with factors a and b:

Factors	$a=1,\ b=1$	$a=1,\ b\neq 1$	$a\neq 1,\ b=1$	$a\neq 1,\ b\neq 1$
Fixed points	every point	x-axis	y-axis	O

Frame 5

Shears



Fixed points of shears with factor a: P

Type	x-shear		y-shear	
Factor	$a\neq 0$	$a=0$	$a\neq 0$	$a=0$
Fixed points	x-axis	every point	y-axis	every point

Frame 6

Activity 1.1 Describing fixed points

Describe the fixed points of the following linear transformations.

(a) $r_{3\pi/2}$ (b) $q_{\pi/2}$ (c) scaling with factors 1 and -1 (d) y-shear with factor 4

Solutions are given on page 52.

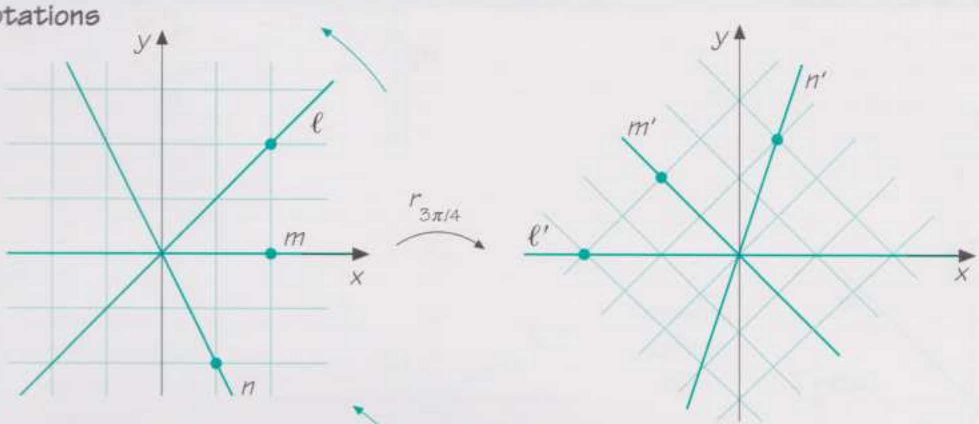
Frame 7

Invariant lines

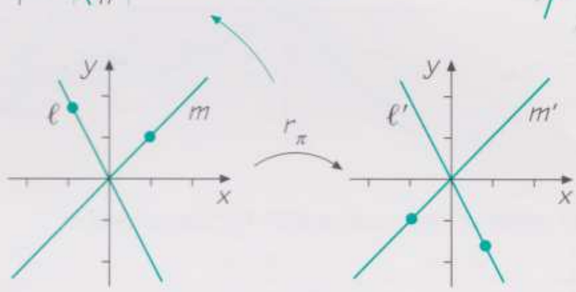
An **invariant line** of a linear transformation is a line that is equal to its image line.

Frame 8

Rotations



no invariant lines



All lines through the origin are invariant lines.

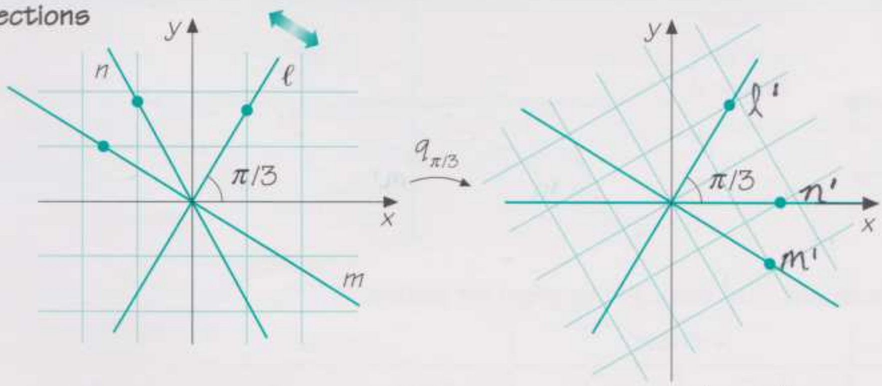
Invariant lines through the origin for the rotation r_θ:

Angle	$\theta \neq k\pi$	$\theta = k\pi$
Invariant lines	none	every line

$k \in \mathbb{Z}$

Frame 9

Reflections



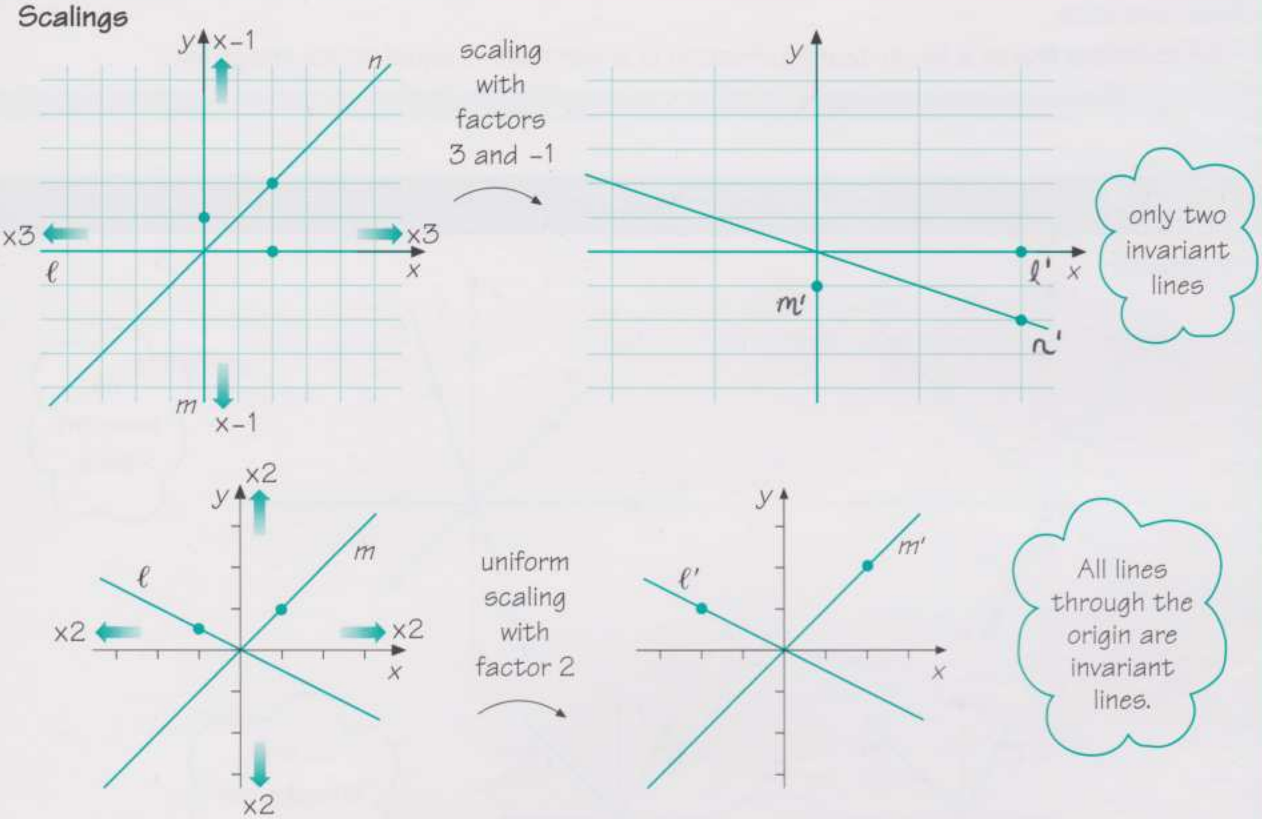
only two invariant lines

Invariant lines through the origin for the reflection q_θ: *l and m*

Two invariant lines – the axis of reflection and the line perpendicular to the axis of reflection.

Frame 10

Scalings

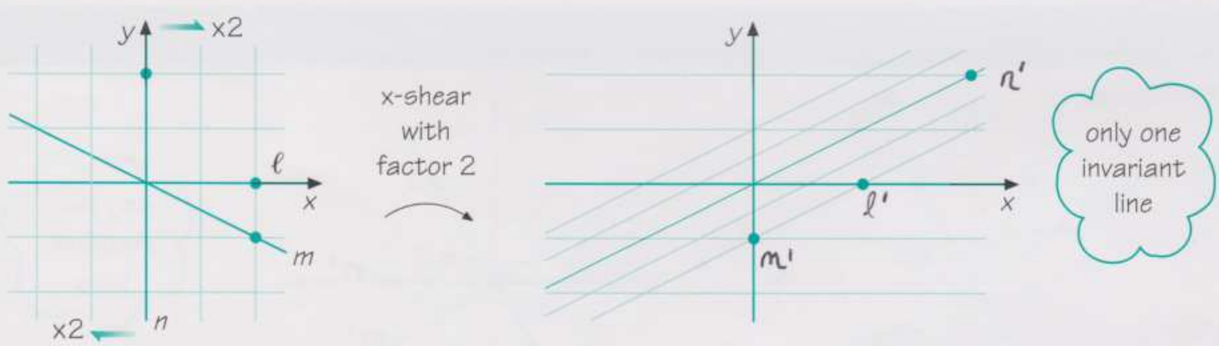


Invariant lines through the origin of the scaling with factors a and b:

Factors	$a \neq b$	$a = b$
Invariant lines	x-axis and y-axis	every line

Frame 11

Shears



Invariant lines through the origin for shears with factor a:

Type	x-shear		y-shear	
Factor	$a \neq 0$	$a = 0$	$a \neq 0$	$a = 0$
Invariant lines	x-axis	every line	y-axis	every line

Activity 1.2 Invariant lines through the origin

Describe the invariant lines through the origin (if there are any) of each of the following linear transformations.

(a) $r_{3\pi/2}$ (b) $q_{\pi/2}$ (c) scaling with factors 1 and -1

(d) y-shear with factor 4

Solutions are given on page 52.

The fixed points and invariant lines through the origin for each of the four types of linear transformation considered in the audio tape are summarised below.

Fixed points

- ◇ Most rotations about the origin move every point except the origin. The only exceptions are rotations through an integer multiple of 2π, in which case *every* point is a fixed point.
- ◇ A reflection in a line through the origin has exactly one line of fixed points: the axis of reflection.
- ◇ Most scalings have no fixed points except the origin. The exceptions occur when one (or both) of the factors is one.
- ◇ Most shears have exactly one line of fixed points: for an x-shear this line is the x-axis, for a y-shear this line is the y-axis. The exceptions are shears with factor zero, for which every point is a fixed point.

Invariant lines through the origin

There may also be other invariant lines not passing through the origin.

- ◇ Most rotations about the origin have no invariant lines. The exceptions are rotations through an integer multiple of π, in which case *every* line through the origin is an invariant line.
- ◇ For any reflection in a line through the origin, the axis of reflection is an invariant line consisting of fixed points. The line (through the origin) perpendicular to the axis of reflection is also an invariant line.
- ◇ A scaling with factors a and b has at least two invariant lines: the coordinate axes. However, if $a = b$ then *every* line through the origin is an invariant line.

Remember that the scale factors a and b are both non-zero.

- ◇ Most shears have just one invariant line through the origin: for an x-shear this line is the x-axis and for a y-shear this line is the y-axis. However, if the shear has factor 0, then every line through the origin is an invariant line.

We have used our geometric understanding of these four types of linear transformation to determine their fixed points and invariant lines through the origin. If we know the matrix which represents one of these linear transformations, then we can check our results using algebra.

For example, reflection in the x-axis is represented by the matrix $\mathbf{Q}_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. From our geometric understanding of reflections, we know that every point on the x-axis (the axis of reflection) is a fixed point. To check this algebraically, we find the image of an arbitrary point on the

See Chapter B2, Subsection 1.3.

x-axis, say $(c, 0)$. We represent the point $(c, 0)$ by the vector $\mathbf{x} = \begin{pmatrix} c \\ 0 \end{pmatrix}$ and show that $\mathbf{Ax} = \mathbf{x}$, where $\mathbf{A} = \mathbf{Q}_0$. In fact,

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} c+0 \\ 0+0 \end{pmatrix} = \begin{pmatrix} c \\ 0 \end{pmatrix},$$

so $(c, 0)$ is a fixed point, as expected. Since $(c, 0)$ is an arbitrary point on the x-axis, we have shown that every point on that axis is a fixed point.

We can define a fixed point of a linear transformation algebraically as follows.

Algebraic definition of a fixed point

A **fixed point** of a linear transformation represented by the matrix $\mathbf{A}$ is a point, represented by the vector $\mathbf{x}$, such that $\mathbf{Ax} = \mathbf{x}$.

Indeed, any line perpendicular to the x-axis is an invariant line.

Similarly, from our geometric understanding of reflections, we know that for the above reflection, the y-axis is an invariant line; that is, that the image of each point on the y-axis lies on the y-axis. We now use algebra to check this fact. Let $(0, c)$ be an arbitrary point on the y-axis. The image of the point $(0, c)$ is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ -c \end{pmatrix}.$$

The point $(0, -c)$ lies on the y-axis. Since we started with an arbitrary point on the y-axis, we have shown that every point on the y-axis has its image on the y-axis. Also, as c varies, this image ranges over the whole y-axis. Thus the y-axis is an invariant line for this reflection.

Activity 1.3 Checking a fixed point and an invariant line

Use the matrix that represents a scaling with factors 1 and 3 to show algebraically that:

(a) the point $(2, 0)$ is a fixed point of this scaling;

(b) the y-axis is an invariant line of this scaling.

Solutions are given on page 52.

Given a line and the matrix representing a linear transformation, we can check whether that line is an invariant line of the linear transformation as in Activity 1.3. In Section 2, we address the problem of determining such invariant lines.

Summary of Section 1

This section has introduced:

- ◇ the ideas of a fixed point and an invariant line of a linear transformation;
- ◇ the determination, by geometric visualisation, of the fixed points and invariant lines through the origin for rotations, reflections, scalings and shears;
- ◇ a matrix method for checking such determinations.

Exercises for Section 1

Exercise 1.1

Describe the fixed points and invariant lines through the origin of each of the following linear transformations.

(a) r_π (b) $q_{\pi/4}$ (c) scaling with factors 2 and -3

(d) x-shear with factor 1

Exercise 1.2

This exercise concerns the x-shear with factor 1.

(a) Use the matrix representing this shear to check that every point on the x-axis is a fixed point of this shear.

(b) Show that the line $y = x$ is *not* an invariant line of this shear. (*Hint*: Find a point on the line $y = x$ whose image does not lie on the line $y = x$.)

2 Eigenvalues and eigenlines

In the previous section, we used our knowledge of the geometric effect of certain types of linear transformation to investigate their fixed points and invariant lines through the origin. Given the matrix which represents one of these linear transformations, you saw how to check such geometric results using algebra. However, for a linear transformation represented by a general 2×2 matrix, you might not have a clear understanding of the geometric effect of the linear transformation on the plane. For such a linear transformation, the algebraic definition of a fixed point, given in Section 1, can be used to identify any fixed points of the linear transformation. Identifying the invariant lines through the origin is not so straightforward, and in this section we develop an algebraic method to do this. Although identifying the fixed points of a linear transformation does give us some information about its geometric effect, identifying the invariant lines through the origin gives us far more information.

Throughout the remainder of this chapter, we often consider points in the plane $\mathbb{R}^2$ and use the vector representation of these points in algebraic manipulations. When we talk about transformations of the plane in terms of matrices, the point (x, y) will be represented by its position vector $\begin{pmatrix} x \\ y \end{pmatrix}$. Generally, the word 'point' is used when discussing the plane, and the word 'vector' is used when discussing the related algebra. But to avoid cumbersome expressions when both contexts are relevant, the words point and vector will often be used interchangeably.

2.1 Finding eigenvalues and eigenlines

As a first step towards an algebraic method for finding the invariant lines through the origin for a linear transformation, we need to look more closely at the images of points on such invariant lines.

Activity 2.1 Finding invariant lines

Consider the linear transformation

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x}, \quad \text{where } \mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}.$$

(a) By finding the image of the arbitrary vector $\begin{pmatrix} c \\ -c \end{pmatrix}$ on the line $y = -x$, show that this line is an invariant line of the linear transformation f.

(b) Find the image under f of each of the following vectors, which lie on the line $y = -x$.

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 3 \\ -3 \end{pmatrix} \quad \begin{pmatrix} -4 \\ 4 \end{pmatrix}$$

Express each image as a scalar multiple of the vector whose image it is.

(c) By finding the image of the arbitrary vector $\begin{pmatrix} c \\ \frac{3}{2}c \end{pmatrix}$ on the line $y = \frac{3}{2}x$, show that this line is an invariant line of the linear transformation f.

(d) Find the image under f of each of the following vectors, which lie on the line $y = \frac{3}{2}x$.

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} \quad \begin{pmatrix} -2 \\ -3 \end{pmatrix}$$

Express each image as a scalar multiple of the vector whose image it is.

Solutions are given on page 52.

In Activity 2.1(a) and (c), you saw that the lines $y = -x$ and $y = \frac{3}{2}x$ are invariant lines of the linear transformation f.

In Activity 2.1(b), you saw that the image of each given vector is -1 times the given vector. For example,

$$f\begin{pmatrix} 3 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = -1\begin{pmatrix} 3 \\ -3 \end{pmatrix}.$$

In general, the image of the vector $\begin{pmatrix} c \\ -c \end{pmatrix}$ is $\begin{pmatrix} -c \\ c \end{pmatrix} = -1\begin{pmatrix} c \\ -c \end{pmatrix}$, as shown in Figure 2.1.

Also, in Activity 2.1(d), you saw that the image of each given vector is 4 times the given vector. For example,

$$f\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} = 4\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix}.$$

In general, the image of the vector $\begin{pmatrix} c \\ \frac{3}{2}c \end{pmatrix}$ is $\begin{pmatrix} 4c \\ 6c \end{pmatrix} = 4\begin{pmatrix} c \\ \frac{3}{2}c \end{pmatrix}$, as shown in Figure 2.1.

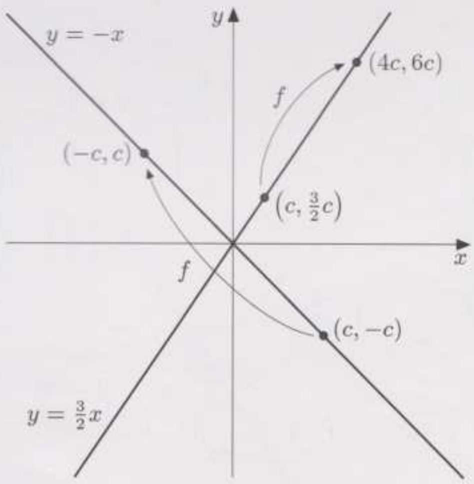


Figure 2.1 The effect of f on points on the invariant lines $y = -x$ and $y = \frac{3}{2}x$

As indicated in Figure 2.1, each point on an invariant line through the origin is mapped to a point which is a fixed scalar multiple of itself. In fact, as is shown below, this property holds for any invariant line through the origin of any linear transformation.

A scalar multiple of a point (x, y) has the form (cx, cy), where $c \in \mathbb{R}$.

Here

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Suppose that the linear transformation represented by the matrix $\mathbf{A}$ has an invariant line through the origin and that (x, y) is a point on this line, represented by the vector $\mathbf{x}$. Then, for some $k \in \mathbb{R}$,

$$\mathbf{A}\mathbf{x} = k\mathbf{x}. \tag{2.1}$$

Now what will be the effect of the transformation on another point on this invariant line? If a point is on the same line through the origin as the point (x, y), then it is a scalar multiple of (x, y), say (cx, cy), where $c \in \mathbb{R}$. So what happens to the point (cx, cy), represented by the vector $c\mathbf{x}$, when we apply the transformation represented by $\mathbf{A}$? We obtain

$$\begin{aligned} \mathbf{A}(c\mathbf{x}) &= c(\mathbf{A}\mathbf{x}), && \text{by the rule for scalar multiplication of matrices,} \\ &= c(k\mathbf{x}), && \text{by equation (2.1),} \\ &= k(c\mathbf{x}). \end{aligned}$$

So every point (cx, cy) on this invariant line is also scalar multiplied by the number k when the transformation is applied. Any line through the origin with this scaling property is called an *eigenline* and the corresponding constant scale factor k is called an *eigenvalue* of the matrix $\mathbf{A}$. Thus we have shown that every invariant line through the origin of a linear transformation is an eigenline of the matrix that represents the transformation. Any vector representing a non-zero point on the eigenline is called an *eigenvector*.

For example, for the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ in Activity 2.1, one eigenline of $\mathbf{A}$ is $y = -x$, with corresponding eigenvalue -1 and another eigenline of $\mathbf{A}$ is $y = \frac{3}{2}x$, with corresponding eigenvalue 4. Each eigenline has many eigenvectors; in fact, any non-zero point on an eigenline corresponds to an eigenvector. Two eigenvectors for the eigenline $y = -x$ are the vectors $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 2 \end{pmatrix}$. Two eigenvectors for the eigenline $y = \frac{3}{2}x$ are the vectors $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} -4 \\ -6 \end{pmatrix}$. These eigenvectors are illustrated in Figure 2.2.

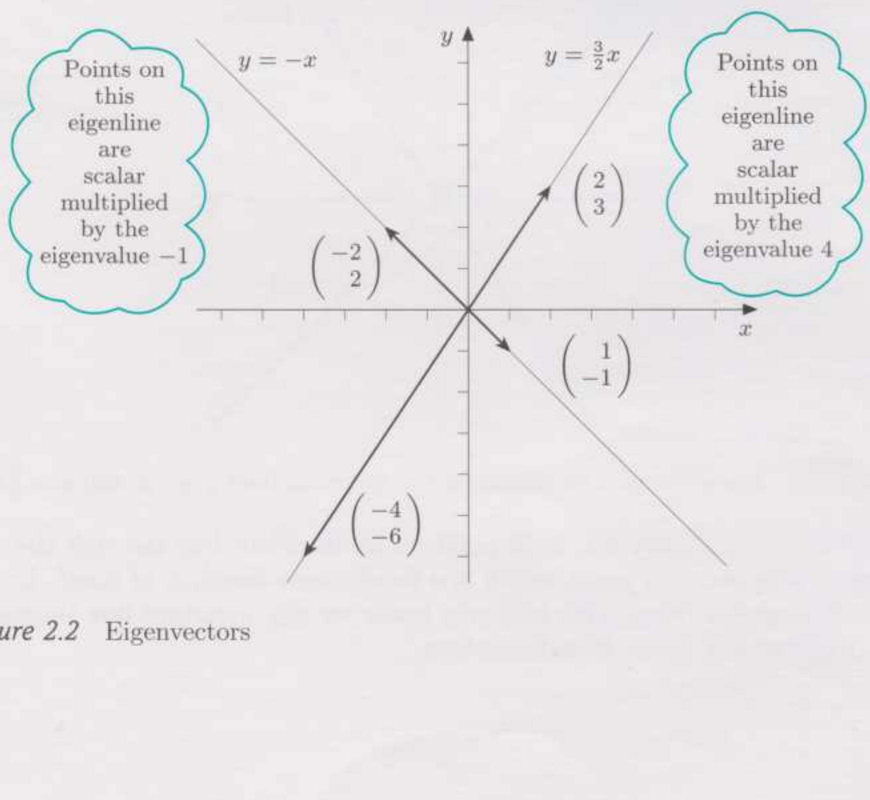


Figure 2.2 Eigenvectors

Here are the formal definitions of the terms introduced above.

Definitions

Let $\mathbf{A}$ be a 2×2 matrix. If $\mathbf{x}$ is a non-zero vector representing a point (x, y) for which there is a real number k such that $\mathbf{Ax} = k\mathbf{x}$, then

- k is called an **eigenvalue** of $\mathbf{A}$ and $\mathbf{x}$ an **eigenvector** of $\mathbf{A}$;
- the line through the origin on which an eigenvector lies is called an **eigenline**;
- $\mathbf{Ax} = k\mathbf{x}$ is called the **eigenvector equation**.

The word 'eigen' is the German adjective meaning 'characteristic' or 'own', so an eigenvalue of a matrix is a number that is characteristic of the matrix.
Many texts use the Greek letters λ (lambda) and μ (mu) to represent eigenvalues.

A number of remarks about these definitions are in order.

1. If k is an eigenvalue of a matrix $\mathbf{A}$ that represents a linear transformation f, then k is often called 'an eigenvalue of f'. Also, the phrases 'eigenvector of f' and 'eigenline of f' are in common use.
2. Not all matrices have eigenvalues. See Activity 2.8, for example.
3. The condition that the eigenvector $\mathbf{x}$ is non-zero is important. If $\mathbf{x}$ is the zero vector, then the equation $\mathbf{Ax} = k\mathbf{x}$ holds for any number k and any 2×2 matrix $\mathbf{A}$. We are interested in picking out the non-zero vectors $\mathbf{x}$ such that $\mathbf{Ax}$ is a scalar multiple of $\mathbf{x}$, and the values of k that appear as 'scale factors'.
4. For each eigenline of $\mathbf{A}$ there is a *corresponding* value of k. For each eigenvalue k, the eigenvectors that satisfy the eigenvector equation are said to *correspond* to that value.
5. It follows from the eigenvector equation, $\mathbf{Ax} = k\mathbf{x}$, that the linear transformation f represented by $\mathbf{A}$ has a certain scaling effect. Figure 2.3 shows this effect on eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ along the two eigenlines, ℓ_1 and ℓ_2, of $\mathbf{A}$. In this figure, the corresponding eigenvalues k_1 and k_2 are taken as positive and negative, respectively.

Note that each eigenline *includes* the origin, but that the vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is *not* an eigenvector.

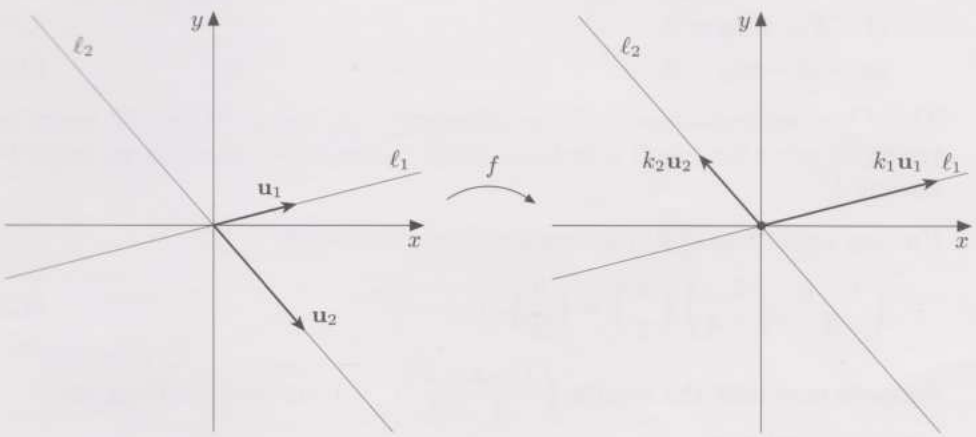


Figure 2.3 Scaling along eigenlines

In general, if (x, y) lies on an eigenline with eigenvalue k, then the image of (x, y) is $|k|$ times further from the origin along the eigenline than (x, y) is. If $k > 0$, then the image lies on the same side of the origin as (x, y). If $k < 0$, then the image lies on the opposite side of the origin from (x, y).

This linear transformation is not invertible because it is not one-one. Hence the matrix representing it is not invertible.

6. You will see later that a matrix can have eigenvalue 0. The linear transformation represented by such a matrix maps all points on the eigenline corresponding to eigenvalue 0 to $(0,0)$. In this case, therefore, the eigenline is *not* an invariant line (since its image is just a point). However, if $\mathbf{A}$ is an invertible matrix that represents an invertible linear transformation f, then the eigenlines of $\mathbf{A}$ and the invariant lines through the origin of f are identical.
7. Equations of the form $\mathbf{Ax} = k\mathbf{x}$ arise in a large variety of mechanical and electrical systems in engineering. For example, the resonant frequencies of bridges, aeroplanes, cars and other mechanical structures are calculated by solving an eigenvector equation. The eigenvector equation also arises in statistical, numerical and other non-engineering situations, such as the modelling of the growth of populations, as you will see in Section 4.

Finding eigenvalues

If we are given the matrix which represents a linear transformation, how can we determine its eigenlines and eigenvalues (if there are any)? We start by looking at an example, before dealing with the general case.

Suppose we want to find the eigenlines and eigenvalues of the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. Let $\mathbf{x}$ be a non-zero vector representing a point (x, y) on an invariant line through the origin. Then $\mathbf{Ax} = k\mathbf{x}$ for some $k \in \mathbb{R}$; that is,

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = k \begin{pmatrix} x \\ y \end{pmatrix}.$$

This matrix equation corresponds to the two simultaneous equations

$$\begin{aligned} x + 2y &= kx, \\ 3x + 2y &= ky; \end{aligned}$$

that is,

$$\begin{aligned} (1-k)x + 2y &= 0, \\ 3x + (2-k)y &= 0. \end{aligned} \tag{2.2}$$

These are two equations in three unknowns x, y and k. We could begin by trying to solve for x and y in terms of k, but in fact it is easier to begin by finding k.

Putting equations (2.2) into matrix form, we obtain

$$\begin{pmatrix} 1-k & 2 \\ 3 & 2-k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.3}$$

The left-hand matrix is $\mathbf{A} - k\mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Suppose now that the matrix $\begin{pmatrix} 1-k & 2 \\ 3 & 2-k \end{pmatrix}$ is invertible. Then we obtain the solution

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1-k & 2 \\ 3 & 2-k \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{aligned}$$

The solution of matrix equations such as equation (2.3) was discussed in MST121 Chapter B2, Subsection 5.3.

However, we are looking for solutions of equation (2.3) in which $\begin{pmatrix} x \\ y \end{pmatrix}$ is non-zero. So the matrix $\begin{pmatrix} 1-k & 2 \\ 3 & 2-k \end{pmatrix}$ must be non-invertible.

Hence the determinant of $\begin{pmatrix} 1-k & 2 \\ 3 & 2-k \end{pmatrix}$ must be zero; that is,

$$(1-k)(2-k) - 6 = 0.$$

The determinant test for non-invertibility of a matrix, based on showing that the determinant of the matrix is 0, was stated in MST121 Chapter B2, Subsection 5.2. Recall that the determinant of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $ad - bc$.

Simplifying this equation, we obtain $k^2 - 3k - 4 = 0$, which can be solved by factorising or by using the quadratic formula to give the eigenvalues $k = 4$ and $k = -1$ (which agree with Activity 2.1).

Once we have found the eigenvalues of a matrix, it is possible to use these values to find the equations of the corresponding eigenlines. We shall return to this task after we have discussed finding eigenvalues in the general case.

By using the method just described to find the eigenvalues for a general 2×2 matrix, we shall find a formula which will allow us to calculate eigenvalues quickly. Suppose that we want to find the eigenvalues of the matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Let $\mathbf{x}$ be a non-zero vector representing a point (x, y) on an eigenline. Then $\mathbf{Ax} = k\mathbf{x}$ for some $k \in \mathbb{R}$; that is,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = k\begin{pmatrix} x \\ y \end{pmatrix}.$$

This matrix equation corresponds to the two simultaneous equations

$$ax + by = kx,$$
$$cx + dy = ky;$$

that is,

$$(a-k)x + by = 0,$$
$$cx + (d-k)y = 0. \qquad (2.4)$$

Putting equations (2.4) into matrix form, we obtain

$$\begin{pmatrix} a-k & b \\ c & d-k \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since we want the vector $\mathbf{x}$ to be non-zero, the matrix

$$\begin{pmatrix} a-k & b \\ c & d-k \end{pmatrix}$$

must be non-invertible; that is, it must have zero determinant. So

$$(a-k)(d-k) - bc = 0.$$

Rearranging this equation, we obtain

$$k^2 - (a+d)k + ad - bc = 0. \qquad (2.5)$$

It may help you to remember this equation if you notice that the coefficient of k is minus the sum of the diagonal entries of $\mathbf{A}$, called the **trace** of $\mathbf{A}$, and that the constant term is the determinant of $\mathbf{A}$.

This quadratic equation is called the **characteristic equation** of $\mathbf{A}$.

Thus, for a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the eigenvalues of $\mathbf{A}$ are the solutions (if any) of the characteristic equation of $\mathbf{A}$. This equation enables us to quickly calculate the eigenvalues of a 2×2 matrix – if there are any. If the characteristic equation has no real solutions, then $\mathbf{A}$ has no eigenvalues.

A handy way to *check* that you have correctly calculated the eigenvalues of $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is to check that the sum of the eigenvalues is equal to the trace $a + d$, the sum of the diagonal elements of the matrix $\mathbf{A}$. For example, you have seen that the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ has eigenvalues -1 and 4. Here, the sum of the eigenvalues is $-1 + 4 = 3$, and the trace is $a + d = 1 + 2 = 3$ as expected.

When finding the eigenvalues of a particular matrix, there is no need to repeat the argument deriving the characteristic equation given above. Just write down the characteristic equation, and solve it by factorising or by using the quadratic formula.

Example 2.1 Finding eigenvalues

Find the eigenvalues (if they exist) of each of the following matrices.

(a) $\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} -2 & 7 \\ -1 & 3 \end{pmatrix}$

Solution

(a) We have $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$, so the characteristic equation is

$$k^2 - (2+4)k + 8 - 3 = 0;$$

that is, $k^2 - 6k + 5 = 0$. This equation can be factorised:

Here the trace is 6 and the determinant is 5, so you can write down

$$k^2 - 6k + 5 = 0$$

directly.

$$(k-5)(k-1) = 0,$$

so the eigenvalues are 1 and 5. (A quick check shows that the sum of the eigenvalues, $1 + 5 = 6$, is equal to the trace, $2 + 4 = 6$.)

(b) We have $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -2 & 7 \\ -1 & 3 \end{pmatrix}$, so the characteristic equation is

$$k^2 - k + 1 = 0.$$

If we try to solve this quadratic equation using the formula, then we obtain

$$k = \tfrac{1}{2}\left(1 \pm \sqrt{(-1)^2 - 4 \times 1}\right).$$

Now $(-1)^2 - 4 \times 1 = -3$ is negative, so this quadratic equation has no real solutions. Hence this matrix has no eigenvalues.

Now it is your turn to find some eigenvalues.

Activity 2.2 Finding eigenvalues

For each of the following matrices, write down the characteristic equation and use it to find the eigenvalues (if they exist).

(a) $\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -3 \\ 5 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} -0.5 & 1.3 \\ 0.2 & 0.6 \end{pmatrix}$

Solutions are given on page 52.

Finding eigenlines

Now that we know how to find the eigenvalues for a 2×2 matrix (when they exist), how can we use these eigenvalues to find the corresponding eigenlines? We start by returning to our initial example, involving the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. We found that this matrix has eigenvalues $k = -1$ and $k = 4$.

To find the equation of the eigenline corresponding to the eigenvalue $k = -1$, we substitute $k = -1$ into the eigenvector equation $\mathbf{Ax} = k\mathbf{x}$. This gives

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -1 \begin{pmatrix} x \\ y \end{pmatrix}.$$

This matrix equation corresponds to the following two equations

$$\begin{aligned} x + 2y &= -x, \\ 3x + 2y &= -y; \end{aligned}$$

that is,

$$\begin{aligned} 2x + 2y &= 0, \\ 3x + 3y &= 0. \end{aligned}$$

These equations *both* reduce to $x + y = 0$. Thus the eigenline corresponding to the eigenvalue $k = -1$ has equation $y = -x$.

We repeat this procedure with the eigenvalue $k = 4$ to find the other eigenline. We have

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 4 \begin{pmatrix} x \\ y \end{pmatrix}.$$

This matrix equation corresponds to the following two equations

$$\begin{aligned} x + 2y &= 4x, \\ 3x + 2y &= 4y; \end{aligned}$$

that is,

$$\begin{aligned} -3x + 2y &= 0, \\ 3x - 2y &= 0. \end{aligned}$$

These equations both reduce to $3x - 2y = 0$. Thus the eigenline corresponding to the eigenvalue $k = 4$ has equation $y = \frac{3}{2}x$.

Notice in this example that, of the pair of equations obtained from the eigenvector equation, one is redundant each time. This always happens (provided that both equations involve both x and y). This fact provides a useful practical check on your working: having found the eigenline using one of the equations, you can check that the other equation is satisfied too. If not, you have made a mistake!

Once we know the equations of the eigenlines, we can choose eigenvectors for each eigenline. An eigenvector for the eigenline $y = -x$ is any non-zero vector $\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}$ for which the corresponding point (u, v) lies on the eigenline. In this case, one eigenvector is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Two other eigenvectors are $\begin{pmatrix} 2 \\ -2 \end{pmatrix}$ and $\begin{pmatrix} -\sqrt{5} \\ \sqrt{5} \end{pmatrix}$.

It is often convenient to have small integer values for the components of the eigenvectors. To choose an eigenvector with integer components, start by writing the equation of the eigenline in the form $y = mx$. Then choose a small non-zero integer value for x such that mx is also an integer.

To choose such a 'nice' eigenvector for the eigenline $y = \frac{3}{2}x$, we need to choose a small non-zero integer value for x such that $\frac{3}{2}x$ is also an integer. In this case, any even integer except 0, will do.

Two other eigenvectors are $\begin{pmatrix} 4 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -8 \\ -12 \end{pmatrix}$.

By choosing $x = 2$, we obtain the eigenvector $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

We usually do try to pick small integer values for the components of eigenvectors, but in fact any vector of the form $\begin{pmatrix} c \\ -c \end{pmatrix}$, where $c \in \mathbb{R}$ and $c \neq 0$, is an eigenvector corresponding to the eigenline $y = -x$. Similarly, any vector of the form $\begin{pmatrix} c \\ \frac{3}{2}c \end{pmatrix}$, where $c \in \mathbb{R}$ and $c \neq 0$, is an eigenvector corresponding to the eigenline $y = \frac{3}{2}x$.

Activity 2.3 Finding eigenlines

For each of the following matrices, use the eigenvalues you found in Activity 2.2 to find the equations of the corresponding eigenlines. For each eigenline, give two eigenvectors.

(a) $\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} -0.5 & 1.3 \\ 0.2 & 0.6 \end{pmatrix}$

Solutions are given on page 53.

Here is a summary of the process for finding eigenvalues, eigenlines and eigenvectors in the form of a strategy.

Strategy

To find the eigenvalues, eigenlines and eigenvectors for the 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1. Solve the characteristic equation $k^2 - (a + d)k + ad - bc = 0$ to find any eigenvalues k. If there are no real solutions to this equation, then there are no eigenvalues.
2. For each eigenvalue k found in Step 1:
   (a) substitute the eigenvalue k into the eigenvector equation $\mathbf{Ax} = k\mathbf{x}$, where $\mathbf{x}$ is the vector $\begin{pmatrix} x \\ y \end{pmatrix}$;
   (b) use the simultaneous equations given by the eigenvector equation to find the equation of the eigenline;
   (c) choose a convenient non-zero vector on the eigenline as an eigenvector.

Remarks

1. As you will see in Subsection 2.2, it is possible to *write down* the eigenvalues of certain types of matrices without having to solve the characteristic equation.
2. If the eigenline has equation $y = mx$, then any non-zero vector of the form $\begin{pmatrix} c \\ mc \end{pmatrix}$ is an eigenvector.
3. You may be aware that when a quadratic equation has no real solutions, it does have solutions that are *complex*. In some contexts, complex eigenvalues are very important. For this course, however, we are interested only in real eigenvalues; so if there are none, then we say that 'the matrix $\mathbf{A}$ has no eigenvalues' (rather than 'the matrix $\mathbf{A}$ has complex eigenvalues'). This corresponds to the definition given on page 17, which requires that an eigenvalue be a real number.

Complex numbers are studied in Chapter D1.

Now try working through the complete process of finding eigenvalues, eigenlines and eigenvectors.

Activity 2.4 Finding eigenvalues, eigenlines and eigenvectors

Let $\mathbf{A}$ be the matrix

$$\begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix}.$$

Find the eigenvalues and eigenlines of $\mathbf{A}$ and, for each eigenvalue, give one eigenvector.

A solution is given on page 53.

2.2 Special cases of eigenvalues

In this subsection, we look first at particular types of matrices for which the eigenvalues are easy to spot. Then we look at matrices which have one eigenvalue equal to zero. Finally, we look at examples of matrices which have no eigenvalues or only one eigenvalue.

Diagonal and triangular matrices

Activity 2.5 Eigenvalues of diagonal matrices

(a) Use the characteristic equation to find the eigenvalues of the diagonal matrix $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$. What do you notice?

(b) Show that the matrix $\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ has eigenvalues a and d.

Diagonal matrices were introduced in Chapter B2, Subsection 2.2.

Solutions are given on page 54.

In Activity 2.5(b), you showed that 2×2 diagonal matrices have their eigenvalues on the leading diagonal (from top left to bottom right).

Triangular matrices were introduced in Chapter B2, Subsection 3.1.

In fact, the same argument applies to a triangular matrix. Indeed, a matrix of the form $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ or $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$ also has the characteristic equation

$$k^2 - (a+d)k + ad = 0,$$

which factorises as $(k-a)(k-d) = 0$. Thus the eigenvalues are again the elements a and d on the leading diagonal.

Presented with the matrix $\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$, we can now say immediately that the eigenvalues are -2 and 2, because this is a diagonal matrix. If we recognised this as the matrix of a scaling, then we could argue geometrically that the eigenlines would be the x- and y-axes. The next activity asks you to verify this algebraically.

Activity 2.6 Eigenvectors and eigenlines of a scaling

Use the eigenvalues -2 and 2 to find the eigenlines of the matrix

$$\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$$

A solution is given on page 54.

Zero eigenvalue

In Chapter B2, Subsection 2.3, you looked at a type of linear transformation called a flattening. In particular, you saw that the linear transformation f represented by the matrix $\mathbf{A} = \begin{pmatrix} 4 & 6 \\ 2 & 3 \end{pmatrix}$ flattens the plane onto the line $x - 2y = 0$; see Figure 2.4.

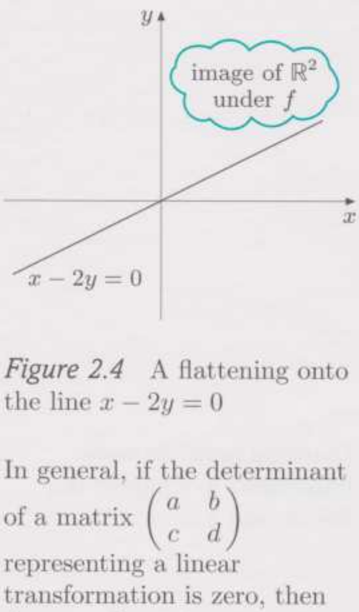


Figure 2.4 A flattening onto the line $x - 2y = 0$

We now investigate further this flattening by finding the eigenvalues and eigenlines of its matrix.

The matrix $\mathbf{A}$ has characteristic equation

$$k^2 - 7k = 0.$$

This factorises as $k(k-7) = 0$, so the eigenvalues of $\mathbf{A}$ are 0 and 7.

For eigenvalue 0, the eigenvector equation,

$$\begin{pmatrix} 4 & 6 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives rise to the two equations $4x + 6y = 0$ and $2x + 3y = 0$. Thus the eigenline corresponding to the eigenvalue 0 is $2x + 3y = 0$.

For eigenvalue 7, the eigenvector equation,

$$\begin{pmatrix} 4 & 6 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 7 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives rise to the two equations $4x + 6y = 7x$ and $2x + 3y = 7y$. Thus the eigenline corresponding to the eigenvalue 7 is $x - 2y = 0$.

Since the eigenline $2x + 3y = 0$ has eigenvalue 0, every point on this line is scalar multiplied by 0; so every point on this line maps to the origin.

In general, if the determinant of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ representing a linear transformation is zero, then the transformation is a flattening and the characteristic equation of the matrix is

$$k^2 - (a+d)k = 0;$$

that is

$$k(k - (a+d)) = 0.$$

You saw in Chapter B2, Subsection 2.3, that this flattening maps every point on a line with equation of the form $2x + 3y = c$ to the point with position vector $c\begin{pmatrix}2\\1\end{pmatrix}$. Thus every point of the plane is mapped to a point on the eigenline $x - 2y = 0$. As a check, notice that if ℓ is the line $2x + 3y = c$, which is parallel to the eigenline $2x + 3y = 0$, then ℓ meets the eigenline $x - 2y = 0$ at the point $(\frac{2}{7}c, \frac{1}{7}c)$. This point is indeed mapped to the point $(2c, c)$ since the eigenvalue of $x - 2y = 0$ is 7, as shown in Figure 2.5.

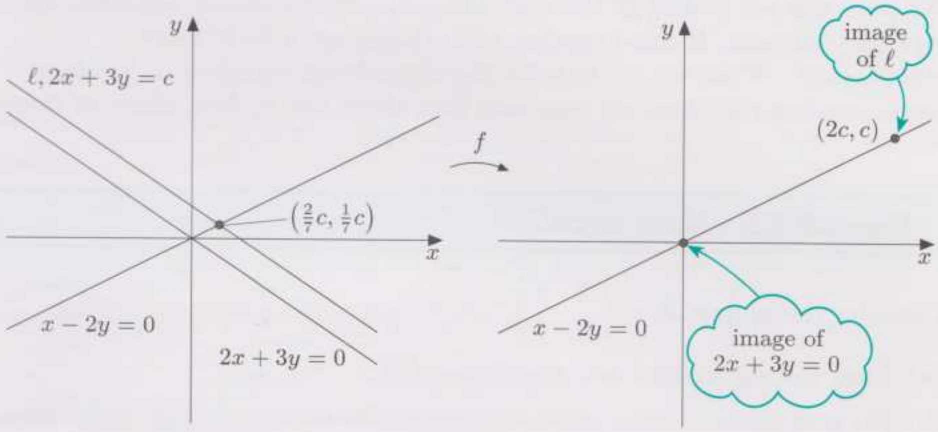


Figure 2.5 A flattening onto the eigenline $x - 2y = 0$, which has eigenvalue 7

In the above discussion, it was pointed out that every point on the eigenline $2x + 3y = 0$ maps to the origin. This is an example of an eigenline which is not an invariant line, as mentioned in Remark 6 on page 18.

The next activity concerns another flattening.

Activity 2.7 Another flattening

The flattening represented by the matrix $\mathbf{A} = \begin{pmatrix}6 & 2\\9 & 3\end{pmatrix}$ collapses the plane onto the line $3x - 2y = 0$.

You saw this in Chapter B2, Activity 2.8.

(a) Find the eigenvalues of $\mathbf{A}$.

(b) Deduce that the flattening maps every point of the eigenline $3x + y = 0$ to the origin and that $3x - 2y = 0$ is also an eigenline of $\mathbf{A}$.

Solutions are given on page 54.

No eigenvalues or one eigenvalue

The characteristic equation of a 2×2 matrix is a quadratic equation, and a quadratic equation may have 0, 1 or 2 (real) solutions. You have seen several examples of matrices which have two distinct eigenvalues, and hence two eigenlines. We now take a closer look at some examples of matrices which have either no eigenvalues or only one eigenvalue. Recall from Section 1 that most rotations r_θ have no invariant lines through the origin; so we would expect a matrix which represents a rotation (through an angle other than a multiple of π) not to have any eigenvalues. The next activity asks you to verify this fact for the rotation through $\frac{1}{2}\pi$.

Activity 2.8 A case of no eigenvalues

If $\theta = \frac{1}{2}\pi$, then

$$\mathbf{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Use the characteristic equation to show that the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, which represents a rotation through an angle of $\frac{1}{2}\pi$, has no eigenvalues.

A solution is given on page 54.

We have not yet looked at the case where the characteristic equation has a repeated solution. If this happens, then things are a little more complicated. When we try to solve the eigenvector equation to locate a corresponding eigenline, we *may* find that there are, in fact, many of them.

Example 2.2 Many eigenlines

Consider the matrix $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, which represents a uniform scaling.

(a) Find the eigenvalues and eigenlines of $\mathbf{A}$.

(b) Do your answers agree with your geometric understanding of the linear transformation represented by this matrix?

Solution

(a) The matrix $\mathbf{A}$ is diagonal, so its eigenvalues are the elements on the leading diagonal. Thus $\mathbf{A}$ has only one eigenvalue, $k = 2$. The eigenvector equation,

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix},$$

gives rise to the two equations $2x = 2x$ and $2y = 2y$. These equations are true for all values of x and y. This implies that every non-zero point (x, y) corresponds to an eigenvector, so *every* line through the origin is an eigenline for this matrix!

(b) We saw in Section 1 that for a uniform scaling, every line through the origin is an invariant line, so the above result agrees with our geometric understanding of uniform scalings.

On the other hand, when the characteristic equation has a repeated solution there may be only a *single* eigenline.

Activity 2.9 A single eigenline

Consider the matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, which represents a y-shear with factor 1.

(a) Find the eigenvalues and eigenlines of $\mathbf{A}$.

(b) Do your answers agree with your geometric understanding of the linear transformation represented by this matrix?

Solutions are given on page 54.

In this course, we consider only 2×2 matrices. The definitions of eigenvalues and eigenvectors can readily be extended to square matrices of larger sizes; but the geometric interpretation of the corresponding transformations is harder to visualise. Also, calculation by hand of eigenvalues and eigenvectors becomes impractical for large matrices. In such cases, a computer can be used.

Summary of Section 2

The main theme of this section was how to find eigenvalues and eigenlines for 2×2 matrices algebraically. In a geometric context, this allowed us to identify any invariant lines through the origin for a linear transformation represented by a 2×2 matrix. This section introduced:

- ◇ eigenvalues, eigenlines and eigenvectors for a 2×2 matrix;
- ◇ examples of matrices which have two eigenvalues and hence two eigenlines;
- ◇ examples of matrices with no eigenvalues and hence no eigenlines;
- ◇ examples of matrices with one eigenvalue, one example having many eigenlines and one example having only one eigenline;
- ◇ the fact that diagonal and triangular matrices have their eigenvalues on the leading diagonal.

Exercises for Section 2

Exercise 2.1

Find the eigenvalues (if they exist) and eigenlines of each of the following matrices. For each eigenline, give one eigenvector.

(a) $\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

Exercise 2.2

For each matrix below, identify which type of basic linear transformation it represents (rotation, reflection, scaling, shear, flattening) and state how many eigenlines it should have, based on your geometric understanding of the linear transformation. Verify your answers by finding the eigenvalues and eigenlines for each of these matrices.

(a) $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

3 Using eigenvalues and eigenlines

In this section, it is shown that any 2×2 matrix $\mathbf{A}$ which has two distinct eigenvalues can be expressed in the form

$$\mathbf{A} = \mathbf{PDP}^{-1}, \tag{3.1}$$

where $\mathbf{D}$ is a diagonal matrix having the eigenvalues of $\mathbf{A}$ on its leading diagonal, and $\mathbf{P}$ is an invertible matrix whose columns are two eigenvectors of $\mathbf{A}$.

Recall that a matrix $\mathbf{P}$ is invertible if there exists an inverse matrix $\mathbf{P}^{-1}$ such that $\mathbf{PP}^{-1} = \mathbf{I} = \mathbf{P}^{-1}\mathbf{P}$, where $\mathbf{I}$ is the identity matrix.

As you will see, being able to write $\mathbf{A}$ in the form $\mathbf{PDP}^{-1}$ enables the nth power of $\mathbf{A}$ to be expressed in the form

$$\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1},$$

which provides a systematic method for calculating powers of $\mathbf{A}$.

Finally we use eigenvalues and eigenvectors to extend our geometric understanding of certain linear transformations.

3.1 Diagonalising a matrix

In this subsection, you will see how to find a diagonal matrix $\mathbf{D}$ and an invertible matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$ for any 2×2 matrix $\mathbf{A}$ which has two distinct eigenvalues.

Consider a general 2×2 matrix $\mathbf{A}$ and suppose that we can write $\mathbf{A}$ as the product of the three matrices $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is an unknown diagonal matrix and $\mathbf{P}$ is an unknown invertible 2×2 matrix. What do the matrices $\mathbf{P}$ and $\mathbf{D}$ look like? Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}.$$

Since $\mathbf{A} = \mathbf{PDP}^{-1}$, we can multiply both sides of this equation on the right by $\mathbf{P}$ to eliminate $\mathbf{P}^{-1}$. We obtain

$$\begin{aligned} \mathbf{AP} &= \mathbf{PDP}^{-1}\mathbf{P} \\ &= \mathbf{PD}(\mathbf{P}^{-1}\mathbf{P}) \\ &= \mathbf{PD}, \quad \text{since } \mathbf{P}^{-1}\mathbf{P} = \mathbf{I}. \end{aligned}$$

The matrices in a product can be grouped in any convenient way provided that the order of the matrices is not changed.

The equation $\mathbf{AP} = \mathbf{PD}$ can be written in the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix} \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix};$$

that is,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix} = \begin{pmatrix} k_1p_1 & k_2p_2 \\ k_1q_1 & k_2q_2 \end{pmatrix}.$$

We now break up this product by considering each column of $\mathbf{P}$ separately. The product of $\mathbf{A}$ with the first column of $\mathbf{P}$ is equal to the first column of the matrix $\mathbf{PD}$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} k_1p_1 \\ k_1q_1 \end{pmatrix};$$

that is,

$$\mathbf{A}\begin{pmatrix} p_1 \\ q_1 \end{pmatrix} = k_1 \begin{pmatrix} p_1 \\ q_1 \end{pmatrix}. \tag{3.2}$$

The product of $\mathbf{A}$ with the second column of $\mathbf{P}$ is equal to the second column of the matrix $\mathbf{PD}$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} p_2 \\ q_2 \end{pmatrix} = \begin{pmatrix} k_2p_2 \\ k_2q_2 \end{pmatrix};$$

that is,

$$\mathbf{A}\begin{pmatrix} p_2 \\ q_2 \end{pmatrix} = k_2\begin{pmatrix} p_2 \\ q_2 \end{pmatrix}. \qquad (3.3)$$

Equations (3.2) and (3.3) are exactly of the form of the eigenvector equation of $\mathbf{A}$! So these equations are satisfied *provided that* k_1 and k_2 are eigenvalues of $\mathbf{A}$, with corresponding eigenvectors $\begin{pmatrix} p_1 \\ q_1 \end{pmatrix}$ and $\begin{pmatrix} p_2 \\ q_2 \end{pmatrix}$.

Thus, to obtain the equation $\mathbf{AP} = \mathbf{PD}$, we can take $\mathbf{D}$ to have the eigenvalues of $\mathbf{A}$ as its diagonal elements and the columns of the matrix $\mathbf{P}$ to be corresponding eigenvectors of $\mathbf{A}$.

To write $\mathbf{A}$ in the form $\mathbf{PDP}^{-1}$, we need to calculate $\mathbf{P}^{-1}$ as well.

Since the columns of $\mathbf{P}$ are vectors in different directions, $\det \mathbf{P}$ is non-zero, so $\mathbf{P}^{-1}$ exists.

Example 3.1 Finding the matrices P *and* D

Express the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$$

in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.

Solution

From the preceding discussion we know that the eigenvalues of $\mathbf{A}$ are the diagonal elements of $\mathbf{D}$ and corresponding eigenvectors of $\mathbf{A}$ are the columns of $\mathbf{P}$. We found the eigenvalues and corresponding eigenvectors of $\mathbf{A}$ in Section 2.

See pages 18 and 21.

◇ The eigenvalue 4 has corresponding eigenvector $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

◇ The eigenvalue -1 has corresponding eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Thus we can choose

$$\mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}.$$

Also, $\det \mathbf{P} = -5$, so $\mathbf{P}^{-1}$ exists and is given by

$$\begin{aligned}\mathbf{P}^{-1} &= \frac{1}{-5}\begin{pmatrix} -1 & -1 \\ -3 & 2 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}.\end{aligned}$$

Thus we have $\mathbf{A} = \mathbf{PDP}^{-1}$.

Comment

You can check the above result by multiplying the matrices

$$\mathbf{PDP}^{-1} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}.$$

When there is a number ($\frac{1}{5}$ here) involved in a matrix product, it is usually simpler to take the number to the front and perform the scalar multiplication last. Also, remember that you can multiply a matrix product of the form $\mathbf{ABC}$ as $\mathbf{A}(\mathbf{BC})$ or as $(\mathbf{AB})\mathbf{C}$. Thus, for example,

$$\begin{aligned} \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix} &= \frac{1}{5} \begin{pmatrix} 8 & -1 \\ 12 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 5 & 10 \\ 15 & 10 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}, \quad \text{as required.} \end{aligned}$$

Note that the *order* in which the eigenvalues are entered in the leading diagonal of $\mathbf{D}$ matters. If we had written down the matrix $\mathbf{D}$ as

$$\begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix},$$

then a matrix $\mathbf{P}$ such that $\mathbf{A} = \mathbf{PDP}^{-1}$ would be $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$. In other words, the columns of the matrix $\mathbf{P}$ have to be interchanged! The order of eigenvalues and eigenvectors must be matched up.

For example, we could have used $\mathbf{P} = \begin{pmatrix} 4 & -2 \\ 6 & 2 \end{pmatrix}$, for which $\mathbf{P}^{-1} = \frac{1}{20} \begin{pmatrix} 2 & 2 \\ -6 & 4 \end{pmatrix}$.

Note that there are many invertible matrices $\mathbf{P}$ which could be used in equation (3.1), since there are many possible eigenvectors for each eigenvalue.

We have now developed the following strategy to express a 2×2 matrix $\mathbf{A}$ with two distinct eigenvalues in the form $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. We call this process **diagonalising** the matrix $\mathbf{A}$.

Strategy

To express a 2×2 matrix $\mathbf{A}$ with two distinct eigenvalues in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.

1. Find the eigenvalues k_1 and k_2 of $\mathbf{A}$.
2. Define the diagonal matrix $\mathbf{D}$ by $\mathbf{D} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$.
3. Find the eigenlines corresponding to the eigenvalues k_1 and k_2.
4. Choose
   - ◇ an eigenvector $\begin{pmatrix} p_1 \\ q_1 \end{pmatrix}$ for the eigenvalue k_1;
   - ◇ an eigenvector $\begin{pmatrix} p_2 \\ q_2 \end{pmatrix}$ for the eigenvalue k_2.
5. Use these eigenvectors to define the matrix $\mathbf{P}$ by $\mathbf{P} = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$.
6. Calculate the matrix $\mathbf{P}^{-1}$.

Then $\mathbf{A} = \mathbf{PDP}^{-1}$.

The columns of $\mathbf{P}$ are eigenvectors in different directions, so $\mathbf{P}$ is invertible.

Remarks

1. When applying this strategy, remember that you must match the order of the eigenvalues on the diagonal of $\mathbf{D}$ and that of the corresponding eigenvectors in the columns of $\mathbf{P}$. It does not matter which order you take, as long as you are consistent.
2. As a convention in this chapter, the eigenvalue of larger magnitude will be written in the first column of $\mathbf{D}$; so if the matrix $\mathbf{A}$ has eigenvalues k_1 and k_2 with $|k_1| > |k_2|$, then $\mathbf{D} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$.
3. There are many choices of eigenvectors for the columns of the matrix $\mathbf{P}$. In order to keep the arithmetic simple, try to choose eigenvectors that have small integer components.
4. It is important to remember that this strategy works only for matrices $\mathbf{A}$ which have two *distinct* eigenvalues, so you cannot apply it to matrices which represent linear transformations such as rotations or shears.

Activity 3.1 Using the strategy

Express the matrix $\mathbf{A} = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}$ in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.

A solution is given on page 55.

Comment

Note that in the solution, we could have swapped the order of the eigenvalues and used the matrices $\mathbf{D}$ and $\mathbf{P}$, where

$$\mathbf{D} = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{P} = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}.$$

We could also have chosen different eigenvectors for the columns of $\mathbf{P}$.

3.2 Matrix powers

Now that we know how to diagonalise a 2×2 matrix having two distinct eigenvalues, we look at why diagonalisation is of use in calculating matrix powers. The following activity illustrates a useful property of diagonal matrices.

Activity 3.2 Powers of a diagonal matrix

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ and $\mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$. Calculate the matrices

$$\mathbf{A}^2, \mathbf{A}^3, \mathbf{D}^2 \text{ and } \mathbf{D}^3.$$

Solutions are given on page 55.

As you will have noticed, calculating powers of a general 2×2 matrix can be tiresome, whereas calculating powers of a diagonal matrix is easy. In general, for a diagonal matrix $\mathbf{D} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$, we have

$$\mathbf{D}^n = \begin{pmatrix} a^n & 0 \\ 0 & d^n \end{pmatrix}.$$

Now suppose that we have a 2×2 matrix $\mathbf{A}$ which can be written in the form $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. Then

Remember not to change the order of the matrices when regrouping.

$$\begin{aligned}
\mathbf{A}^2 &= \mathbf{AA} \\
&= (\mathbf{PDP}^{-1})(\mathbf{PDP}^{-1}), \quad \text{since } \mathbf{A} = \mathbf{PDP}^{-1}, \\
&= (\mathbf{PD})(\mathbf{P}^{-1}\mathbf{P})(\mathbf{DP}^{-1}), \quad \text{regrouping the matrices,} \\
&= (\mathbf{PD})\mathbf{I}(\mathbf{DP}^{-1}), \quad \text{since } \mathbf{P}^{-1}\mathbf{P} = \mathbf{I}, \\
&= \mathbf{PD}^2\mathbf{P}^{-1}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbf{A}^3 &= \mathbf{AA}^2 \\
&= (\mathbf{PDP}^{-1})(\mathbf{PD}^2\mathbf{P}^{-1}), \quad \text{from above result for } \mathbf{A}^2, \\
&= (\mathbf{PD})(\mathbf{P}^{-1}\mathbf{P})(\mathbf{D}^2\mathbf{P}^{-1}) \\
&= (\mathbf{PD})\mathbf{I}(\mathbf{D}^2\mathbf{P}^{-1}) \\
&= \mathbf{PD}^3\mathbf{P}^{-1}.
\end{aligned}$$

In general, we have the following result.

Calculating powers

If a 2×2 matrix $\mathbf{A}$ can be written in the form

$$\mathbf{A} = \mathbf{PDP}^{-1}, \quad \text{where } \mathbf{D} \text{ is a diagonal } 2 \times 2 \text{ matrix,}$$

then

$$\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}, \quad \text{for } n = 1, 2, 3, \ldots.$$

Writing a matrix $\mathbf{A}$ in the form $\mathbf{PDP}^{-1}$ where $\mathbf{D}$ is a diagonal matrix greatly simplifies the calculation of powers of $\mathbf{A}$. To calculate $\mathbf{A}^n$, we first calculate $\mathbf{D}^n$, which is straightforward since $\mathbf{D}$ is a diagonal matrix, and then multiply the three matrices $\mathbf{P}$, $\mathbf{D}^n$ and $\mathbf{P}^{-1}$.

Example 3.2 Matrix powers

As we saw in Example 3.1, the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ can be written in the form $\mathbf{PDP}^{-1}$ where

$$\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbf{P}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}.$$

Use this information to calculate $\mathbf{A}^6$, without multiplying powers of $\mathbf{A}$ directly.

Solution

First we calculate $\mathbf{D}^6$:

$$\mathbf{D}^6 = \begin{pmatrix} 4^6 & 0 \\ 0 & (-1)^6 \end{pmatrix}$$
$$= \begin{pmatrix} 4096 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then we calculate $\mathbf{A}^6$ using the fact that $\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$:

$$\mathbf{A}^6 = \mathbf{P}\mathbf{D}^6\mathbf{P}^{-1}$$
$$= \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 4096 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$
$$= \frac{1}{5} \begin{pmatrix} 8192 & 1 \\ 12\,288 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$
$$= \frac{1}{5} \begin{pmatrix} 8195 & 8190 \\ 12\,285 & 12\,290 \end{pmatrix}$$
$$= \begin{pmatrix} 1639 & 1638 \\ 2457 & 2458 \end{pmatrix}.$$

Now it is your turn to calculate a matrix power.

Activity 3.3 Matrix powers

Consider the matrices

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix}, \mathbf{P} = \begin{pmatrix} -1 & 1 \\ 1 & 4 \end{pmatrix} \text{ and } \mathbf{D} = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}.$$

(a) Calculate the matrix $\mathbf{P}^{-1}$.

(b) Verify that

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}.$$

(c) Hence calculate $\mathbf{A}^4$, without multiplying powers of $\mathbf{A}$ directly.

Solutions are given on page 55.

When calculating a matrix power in this way, it may seem that having first to calculate the matrices $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{P}^{-1}$ is a lengthy preliminary procedure. There are two reasons why diagonalisation is a useful way to calculate matrix powers by hand: speed and accuracy.

If n is a large number, then calculating $\mathbf{D}^n$ is much faster than calculating $\mathbf{A}^n$. Admittedly, diagonalisation also involves having to first calculate the matrices $\mathbf{P}$, $\mathbf{P}^{-1}$ and $\mathbf{D}$, as well as multiplying three matrices together at the end, but if n is a large number, then diagonalisation is still a much quicker way of calculating matrix powers than direct computation.

For example, with $\mathbf{A}$, $\mathbf{P}$ and $\mathbf{D}$ as in Example 3.2,

$$\mathbf{D}^{10} = \begin{pmatrix} 4^{10} & 0 \\ 0 & (-1)^{10} \end{pmatrix} = \begin{pmatrix} 1\,048\,576 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then, using exactly the same form of calculation as for $\mathbf{A}^6$ in Example 3.2, we obtain

$$\mathbf{A}^{10} = \mathbf{P}\mathbf{D}^{10}\mathbf{P}^{-1} = \begin{pmatrix} 419\,431 & 419\,430 \\ 629\,145 & 629\,146 \end{pmatrix}.$$

Calculating $\mathbf{A}^{10}$ directly would involve many more multiplications and additions, so it offers more potential for making arithmetical errors.

Another point to consider here is accuracy. Every time a calculation is carried out by a calculator as an adjunct to hand calculation, the accuracy of the answer is limited by the number of digits stored in memory. The more calculations are done, the more errors creep in due to rounding. Calculation of matrix powers using diagonalisation involves fewer calculations and hence gives more accurate answers.

3.3 Using eigenvalues and eigenlines geometrically

In Chapter B2, you have seen how a linear transformation may be interpreted geometrically by sketching the image of the unit grid. Another grid, based on eigenlines, provides an alternative way to interpret a linear transformation.

Here we consider a linear transformation f represented by a 2×2 matrix $\mathbf{A}$ that has two distinct non-zero eigenvalues. We already know that the image of a point on an eigenline of $\mathbf{A}$ lies on that eigenline, being scaled by the corresponding eigenvalue. But what happens to points that are not on an eigenline?

For example, consider the linear transformation f represented by the matrix $\mathbf{A} = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}$, which, as we have seen in Activity 3.1, has eigenvalues $k_1 = 3$ and $k_2 = -2$, and corresponding eigenlines ℓ_1 and ℓ_2 with equations $y = \frac{1}{3}x$ and $y = -\frac{1}{2}x$, respectively.

On the left of Figure 3.1 the eigenlines of $\mathbf{A}$ are shown, in the domain of f. Let P be a point in the domain of f that does not lie on an eigenline. The following construction shows what happens to P under f.

Draw a parallelogram, based at the origin, by joining P to each eigenline by a line segment parallel to the other eigenline, as shown. Let the vertices of this parallelogram be O, A (on ℓ_2), B (on ℓ_1) and P, as shown. Each of the sides OA and OB of this parallelogram $OAPB$ is scaled under f along the eigenline on which it lies by the corresponding eigenvalue. Since linear transformations preserve parallelism, the image of $OAPB$ is a suitably scaled parallelogram. Let A' and B' denote the images of A and B, respectively. To construct this image parallelogram, mark A' on the line ℓ_2 twice as far from O as A but on the opposite side of line ℓ_1 to A, and mark B' on the line ℓ_1 three times as far from O as B and on the same side of line ℓ_2 as B, as shown on the right of the figure. (In terms of eigenvectors, $\overrightarrow{OA'} = -2\overrightarrow{OA}$ and $\overrightarrow{OB'} = 3\overrightarrow{OB}$.) Now draw the parallelogram with sides OA' and OB'. The fourth vertex of this parallelogram P' is the image of P.

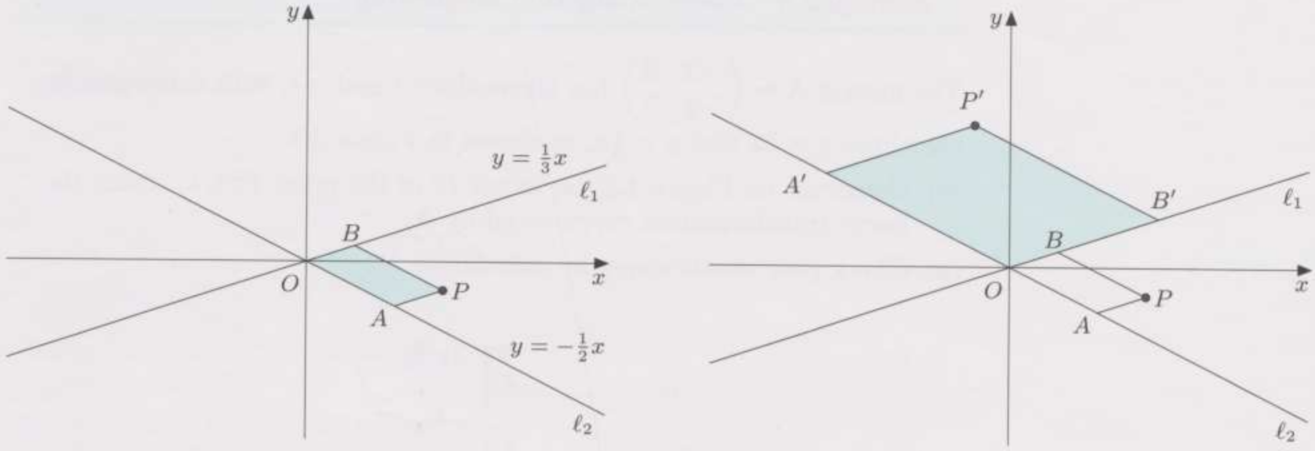


Figure 3.1 Constructing P'

We have located P' by making use of the eigenvalues and eigenlines of $\mathbf{A}$. The effect of f on P can be described as follows: P is scaled by the factor 3 in the direction of the eigenline ℓ_1 and by the factor -2 in the direction of the eigenline ℓ_2. In this sense, the linear transformation f can be described as a *generalised scaling*. A linear transformation represented by a 2×2 matrix that has two distinct non-zero eigenvalues (and hence two distinct eigenlines) is called a **generalised scaling**.

The above construction applies to any point P. In particular, consider P to be such that OA and OB are of unit length. Then the parallelogram $OAPB$ can be thought of as a 'unit parallelogram', which can be used as the basis of a 'unit grid' of parallel lines, each of which is parallel to an eigenline, as shown on the left in Figure 3.2. The image of this grid under f is shown on the right: each point of the original unit grid and each parallelogram has undergone the same generalised scaling.

If P lies on the eigenline ℓ_1, say, then the construction yields the line segment OP', along ℓ_1. The scaling in the direction of ℓ_2 has no effect on P.

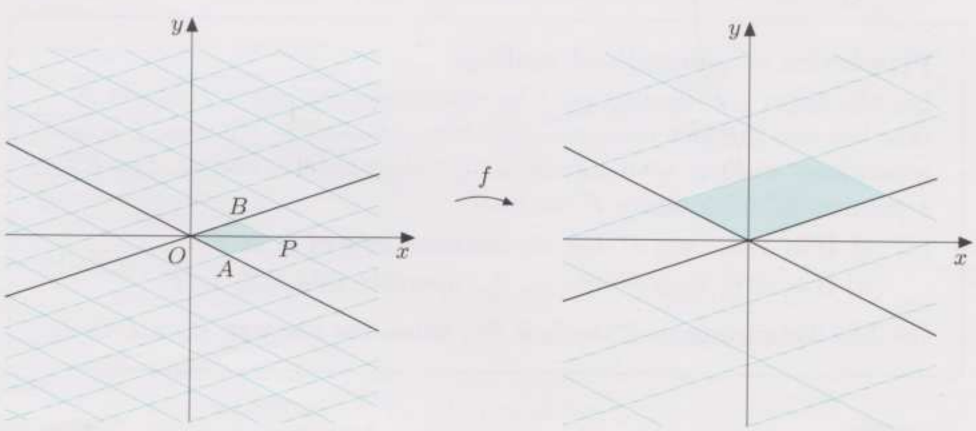


Figure 3.2 A grid of lines parallel to the eigenlines of $\mathbf{A}$, and its image

In the next activity, you are asked to apply the above construction.

Activity 3.4 Constructing and confirming

The matrix $\mathbf{A} = \begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix}$ has eigenvalues 4 and -1, with corresponding eigenlines $y = 2x$ and $y = \frac{1}{3}x$, as shown in Figure 3.3.

(a) Construct on Figure 3.3 the image P' of the point $P(1,1)$ under the linear transformation represented by $\mathbf{A}$.

(b) Check your construction by *calculating* P'.

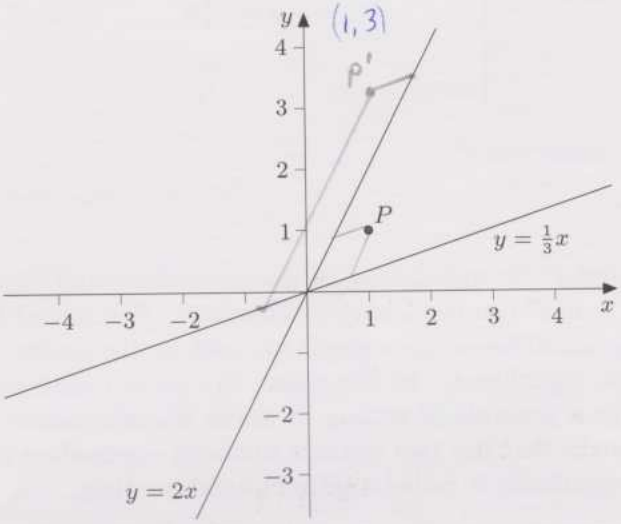


Figure 3.3 Locating P'

Solutions are given on page 55.

This subsection concludes with some properties of generalised scalings, and an activity in which you will use the properties.

Zero eigenvalues are excluded because they correspond to flattenings.

If P lies on ℓ_2, then P' is also on ℓ_2. So there is a sense in which P' lies on the same side of ℓ_2 as P.

Properties of generalised scalings

Let the linear transformation f be represented by a 2×2 matrix $\mathbf{A}$ that has two distinct non-zero eigenvalues k_1 and k_2 (that is, f is a generalised scaling) with corresponding eigenlines ℓ_1 and ℓ_2. Let P be a point of $\mathbb{R}^2$ with image P' under f.

(a) (i) If $k_1 > 0$, then P' lies on the same side of ℓ_2 as P.

(ii) If $k_1 < 0$, then P' lies on the opposite side of ℓ_2 to P.

(b) The distance from P' to ℓ_2 is $|k_1|$ times the distance from P to ℓ_2.

Remarks

1. Corresponding properties are obtained if k_1 and ℓ_2 are replaced by k_2 and ℓ_1.
2. In property (b), the distance from a point to the line ℓ_2 may be taken to be the perpendicular distance from the point to the line, or to be the distance in the direction of the line ℓ_1 from the point to the line ℓ_2. See Figure 3.4.

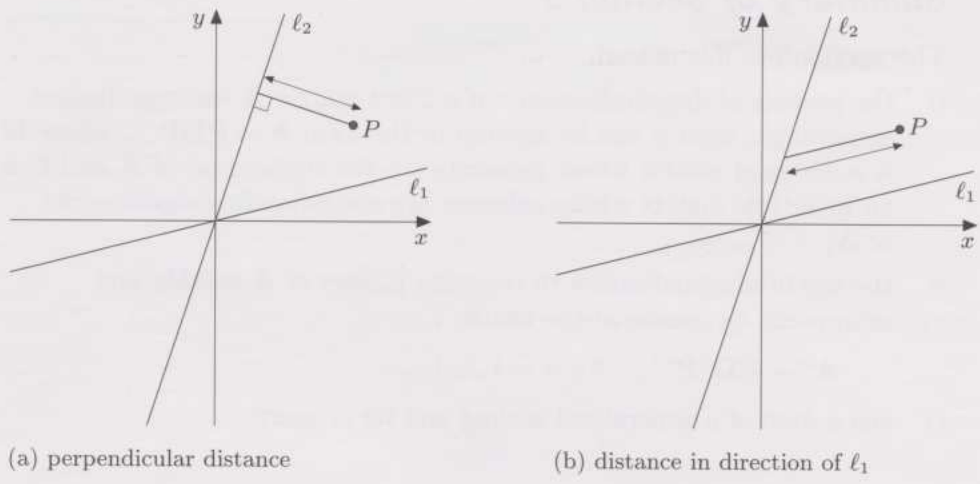


Figure 3.4 Distance from a point to ℓ_2

3. If $|k_1| > 1$, then P' is further from ℓ_2 than P is. See Figure 3.5(a). If $|k_1| < 1$, then P is further from ℓ_2 than P' is. See Figure 3.5(b).

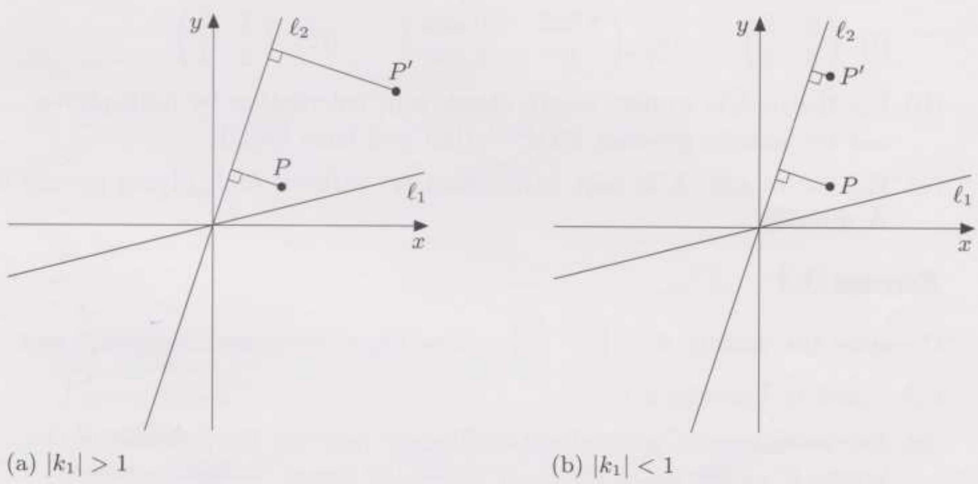


Figure 3.5 Comparing (perpendicular) distances from ℓ_2

Activity 3.5 Locating images

This activity concerns the same matrix and linear transformation as in Activity 3.4.

(a) Use properties of generalised scalings to *describe* the location of

(i) the image P' of the point P shown in Figure 3.3;

(ii) the image P'' of the point P'.

(b) Calculate P''.

Solutions are given on page 56.

Summary of Section 3

This section has introduced:

- the process of diagonalisation – if a 2×2 matrix $\mathbf{A}$ has two distinct eigenvalues, then it can be written in the form $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix whose elements are the eigenvalues of $\mathbf{A}$ and $\mathbf{P}$ is an invertible matrix whose columns are corresponding eigenvectors of $\mathbf{A}$;
- the use of diagonalisation to compute powers of $\mathbf{A}$ quickly and accurately, by means of the result

  $$\mathbf{A}^n = \mathbf{PD}^n\mathbf{P}^{-1}, \quad \text{for } n = 1, 2, 3, \ldots;$$

- the notion of a generalised scaling and its properties.

Exercises for Section 3

Exercise 3.1

(a) Express each of the following matrices in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. (You can use the solution to Exercise 2.1 here.)

(i) $\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix}$ (iii) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

(b) For the matrix in part (a)(i), check your calculation by multiplying out the matrix product $\mathbf{PDP}^{-1}$ that you have found.

(c) For the matrix $\mathbf{A}$ in part (a)(i), find $\mathbf{A}^5$ without multiplying powers of $\mathbf{A}$ directly.

Exercise 3.2

Consider the matrix $\mathbf{A} = \begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$, whose eigenvalues and eigenlines you calculated in Exercise 2.1(a).

(a) Use properties of generalised scalings to *describe* the location of the image P' of the point $P(1, -2)$ under the linear transformation represented by $\mathbf{A}$.

(b) Calculate P'. Draw a sketch to show that the locations of P and P' match your description in part (a).

4 Iterating linear transformations

So far in this chapter, we have looked at the effect of a single application of a linear transformation f. But what happens to the points of $\mathbb{R}^2$ if we repeatedly apply f; that is, if we apply the process of *iteration*?

In Chapter B1, you saw iteration sequences in the context of real functions. Here we look at iteration sequences for linear transformations. Given a linear transformation represented by a matrix $\mathbf{A}$ and an initial point (x_0, y_0), represented by the vector $\mathbf{x}_0$, the **iteration sequence** generated by $\mathbf{A}$ with initial point (x_0, y_0) is the sequence $\mathbf{x}_n$ of points represented by the vectors

$$\mathbf{x}_0, \quad \mathbf{x}_1 = \mathbf{A}\mathbf{x}_0, \quad \mathbf{x}_2 = \mathbf{A}\mathbf{x}_1, \quad \mathbf{x}_3 = \mathbf{A}\mathbf{x}_2, \quad \ldots .$$

$\mathbf{x}_2 = \mathbf{A}^2\mathbf{x}_0$,
$\mathbf{x}_3 = \mathbf{A}^3\mathbf{x}_0, \ldots .$

This is the sequence of points in $\mathbb{R}^2$ obtained by repeated application of the linear transformation represented by the matrix $\mathbf{A}$, starting with the initial point (x_0, y_0). For example, if the matrix $\mathbf{A}$ represents the rotation r_θ and $\mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, then the sequence $\mathbf{x}_n$ appears as in Figure 4.1.

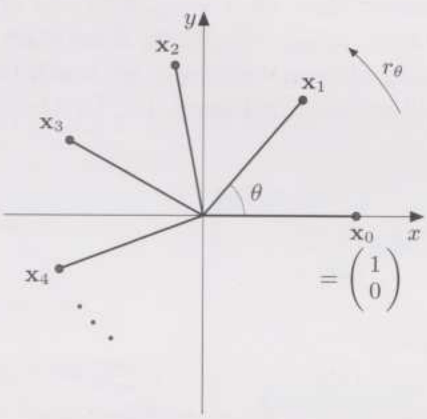


Figure 4.1 The sequence $\mathbf{x}_n$

Given an initial point represented by the vector $\mathbf{x}_0$, the corresponding iteration sequence generated by the matrix $\mathbf{A}$ has recurrence relation

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n \quad (n = 0, 1, 2, \ldots).$$

This sequence can be expressed in closed form as

$$\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0 \quad (n = 1, 2, 3, \ldots).$$

In this section we consider the behaviour of iteration sequences generated by matrices which have two distinct eigenvalues. We first look at such iteration sequences in which the initial point is on one of the eigenlines of $\mathbf{A}$.

Activity 4.1 Iteration of a point on an eigenline

Consider the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}.$$

In particular, see page 21.

Earlier, you saw that $\mathbf{A}$ has eigenvalues 4 and -1 with corresponding eigenlines $y = \frac{3}{2}x$ and $y = -x$. Let $\mathbf{x}_0$ be the vector representing the point $(2, 3)$, which lies on the eigenline $y = \frac{3}{2}x$. Calculate the first three points of the iteration sequence generated by $\mathbf{A}$ with initial point $(2, 3)$.

A solution is given on page 56.

Now consider a general 2×2 matrix $\mathbf{A}$ having two distinct eigenvalues, k_1 and k_2, with corresponding eigenlines ℓ_1 and ℓ_2, respectively. If we start with a point (x_0, y_0) on the eigenline ℓ_1 and apply the linear transformation f represented by $\mathbf{A}$, then the resulting point is (k_1x_0, k_1y_0), which is also on the eigenline ℓ_1. If we apply the same linear transformation to the point (k_1x_0, k_1y_0), then we obtain the point $(k_1^2x_0, k_1^2y_0)$. If we continue iterating this linear transformation, then after n applications of the transformation we obtain the point $(k_1^nx_0, k_1^ny_0)$. Similarly, if we start with a point (x_0', y_0') on the eigenline ℓ_2, then after n iterations, we reach the point $(k_2^nx_0', k_2^ny_0')$. These sequences are illustrated in Figure 4.2, for the case $k_1 > 1$, $k_2 < -1$.

f is a generalised scaling.

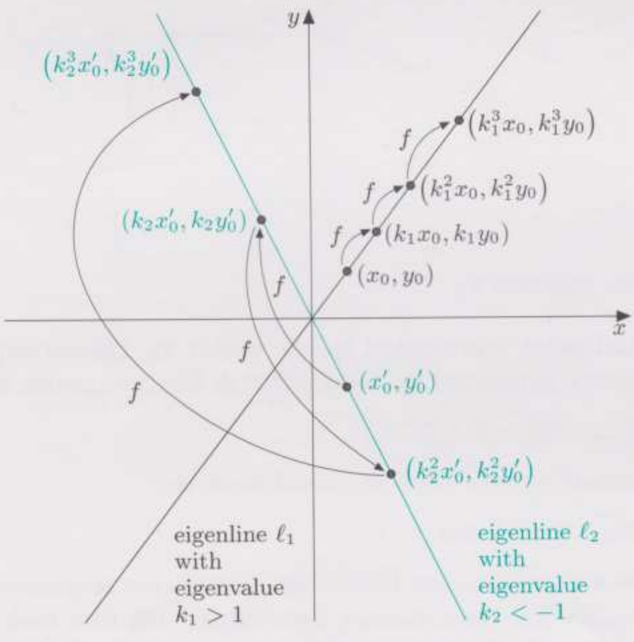


Figure 4.2 Iteration sequences on eigenlines

The table below describes the long-term behaviour of iteration sequences $\mathbf{x}_n$ generated by the matrix $\mathbf{A}$ of a generalised scaling. Here the non-zero initial point $\mathbf{x}_0$ lies on an eigenline ℓ corresponding to a (non-zero) eigenvalue k of $\mathbf{A}$. The table is based on the properties of generalised scalings given in Subsection 3.3.

k	Long-term behaviour of $\mathbf{x}_n$ on ℓ
$k > 1$	$\mathbf{x}_n$ moves away from $(0,0)$, on the same half of ℓ as $\mathbf{x}_0$
$k = 1$	$\mathbf{x}_n = \mathbf{x}_0$, for $n = 0, 1, 2, \ldots$ (a constant sequence)
$0 < k < 1$	$\mathbf{x}_n$ moves towards $(0,0)$, on the same half of ℓ as $\mathbf{x}_0$
$-1 < k < 0$	$\mathbf{x}_n$ moves towards $(0,0)$, alternating between the halves of ℓ
$k = -1$	$\mathbf{x}_n$ alternates between $\pm\mathbf{x}_0$
$k < -1$	$\mathbf{x}_n$ moves away from $(0,0)$, alternating between the halves of ℓ

A line through the origin is divided into two halves by the origin.

In the next activity you will use this table.

Activity 4.2 Long-term behaviour

For the matrix $\mathbf{A}$ of Activity 4.1, use the table above to describe the long-term behaviour of the iteration sequences generated by $\mathbf{A}$ with the following initial points.

(a) $(2,3)$ (b) $(2,-2)$

Solutions are given on page 56.

If we start with an initial point on an eigenline, then the corresponding iteration sequence consists only of points which are on that eigenline. But what can we say about an iteration sequence whose initial point is *not* on an eigenline? Here is an example.

Example 4.1 Iteration of a point not on an eigenline

Consider again the matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ and the point $(2,1)$ represented by the vector $\mathbf{x}_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Note that the point $(2,1)$ satisfies neither of the eigenline equations $y = -x$ and $y = \frac{3}{2}x$.

(a) Calculate the first four points in the iteration sequence

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n \quad (n = 0, 1, 2, \ldots)$$

with initial point $(2,1)$.

(b) Find a formula in terms of n for the vector $\mathbf{x}_n$ which represents the $(n+1)$th point in this iteration sequence. Use that formula to calculate the 10th point, represented by the vector $\mathbf{x}_9$.

Solution

(a) The first point in this iteration sequence is $(2, 1)$.

The second point is represented by the vector

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}.$$

The third point is represented by the vector

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 20 \\ 28 \end{pmatrix}.$$

The fourth point is represented by the vector

$$\mathbf{x}_3 = \mathbf{A}\mathbf{x}_2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 20 \\ 28 \end{pmatrix} = \begin{pmatrix} 76 \\ 116 \end{pmatrix}.$$

Thus the first four points in this iteration sequence are

$$(2,1), (4,8), (20,28), (76,116);$$

see Figure 4.3.

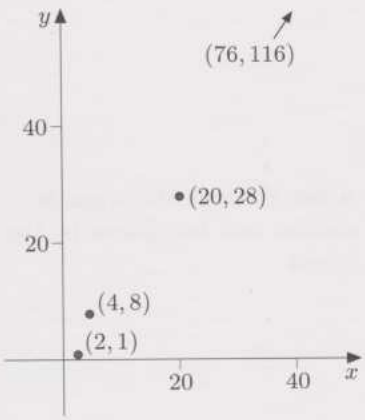


Figure 4.3 The first four points

(b) To find a formula in terms of n for $\mathbf{x}_n$, we use the closed form $\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0$ and the fact that $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, where

We know these matrices from Example 3.2, in which $\mathbf{A}^6$ was calculated.

$$\mathbf{P} = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbf{P}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}.$$

We first calculate a formula for $\mathbf{A}^n$:

In these matrices, to save space, products like 2×4^n have been written as $2(4)^n$.

$$\begin{aligned} \mathbf{A}^n &= \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1} \\ &= \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 4^n & 0 \\ 0 & (-1)^n \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 2(4)^n & (-1)^n \\ 3(4)^n & -(-1)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 2(4)^n + 3(-1)^n & 2(4)^n - 2(-1)^n \\ 3(4)^n - 3(-1)^n & 3(4)^n + 2(-1)^n \end{pmatrix}. \end{aligned}$$

Now we calculate $\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0$:

$$\begin{aligned} \mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0 &= \frac{1}{5}\begin{pmatrix} 2(4)^n + 3(-1)^n & 2(4)^n - 2(-1)^n \\ 3(4)^n - 3(-1)^n & 3(4)^n + 2(-1)^n \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 4(4)^n + 6(-1)^n + 2(4)^n - 2(-1)^n \\ 6(4)^n - 6(-1)^n + 3(4)^n + 2(-1)^n \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 6(4)^n + 4(-1)^n \\ 9(4)^n - 4(-1)^n \end{pmatrix}. \end{aligned}$$

To find the 10th point in this iteration sequence, we substitute $n = 9$ into the expression for $\mathbf{x}_n$:

Remember that the $(n+1)$th point of this sequence is represented by the vector $\mathbf{x}_n$.

$$\begin{aligned} \mathbf{x}_9 &= \frac{1}{5}\begin{pmatrix} 6(4)^9 + 4(-1)^9 \\ 9(4)^9 - 4(-1)^9 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 1\,572\,860 \\ 2\,359\,300 \end{pmatrix} \\ &= \begin{pmatrix} 314\,572 \\ 471\,860 \end{pmatrix}. \end{aligned}$$

Comment

A quicker way to calculate $\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0$ is as follows.

$$\begin{aligned}
\mathbf{A}^n\mathbf{x}_0 &= \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}\mathbf{x}_0 \\
&= (\mathbf{P}\mathbf{D}^n)(\mathbf{P}^{-1}\mathbf{x}_0) \\
&= \frac{1}{5}\left[\begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}\begin{pmatrix} 4^n & 0 \\ 0 & (-1)^n \end{pmatrix}\right]\left[\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] \\
&= \frac{1}{5}\begin{pmatrix} 2(4)^n & (-1)^n \\ 3(4)^n & -(-1)^n \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} \\
&= \frac{1}{5}\begin{pmatrix} 6(4)^n + 4(-1)^n \\ 9(4)^n - 4(-1)^n \end{pmatrix}.
\end{aligned}$$

A disadvantage is that $\mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}$ is not calculated explicitly, and so is not available for future use; see Activity 4.3.

From Figure 4.3, it appears that the points of this iteration sequence are moving up and to the right, away from the origin, in roughly a straight line. The ratio y_n/x_n for each of the calculated points (x_n, y_n) of this iteration sequence gives us an indication of the direction in which the iteration sequence is moving.

Point (x_n, y_n)	$(2,1)$	$(4,8)$	$(20,28)$	$(76,116)$	$\dots$	$(314\,572, 471\,860)$
Ratio y_n/x_n	0.5	2	1.4	1.526 316	$\dots$	1.500 006

It appears that the ratio y_n/x_n is tending towards the value 1.5, so we might conjecture that in the long term, the iteration sequence is moving up and to the right, away from the origin, in the direction of the straight line through the origin with gradient 1.5, that is, the eigenline $y = \frac{3}{2}x$.

Is this behaviour repeated if the initial point is changed to another point not on either of the eigenlines of $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$? In the next activity, you are asked to investigate this question.

Activity 4.3 A different initial point

Consider the matrix $\mathbf{A}$ from Example 4.1, and the initial point $(-3, -1)$.

(a) Calculate the first four points in the iteration sequence

$$\mathbf{x}'_{n+1} = \mathbf{A}\mathbf{x}'_n \quad (n = 0, 1, 2, \dots)$$

with $\mathbf{x}'_0 = \begin{pmatrix} -3 \\ -1 \end{pmatrix}$.

(b) Use the matrix $\mathbf{A}^n$ calculated in Example 4.1 to find a formula in terms of n for the vector $\mathbf{x}'_n$, which represents the $(n+1)$th point in this iteration sequence. Use that formula to calculate the 12th point in this iteration sequence.

(c) Calculate the ratio y'_n/x'_n for each of the points (x'_n, y'_n) found in parts (a) and (b). Do these ratios appear to tend to a particular value?

Solutions are given on page 56.

The first four points for each of the two iteration sequences $\mathbf{x}_n$ and $\mathbf{x}'_n$, calculated above, are plotted in Figure 4.4.

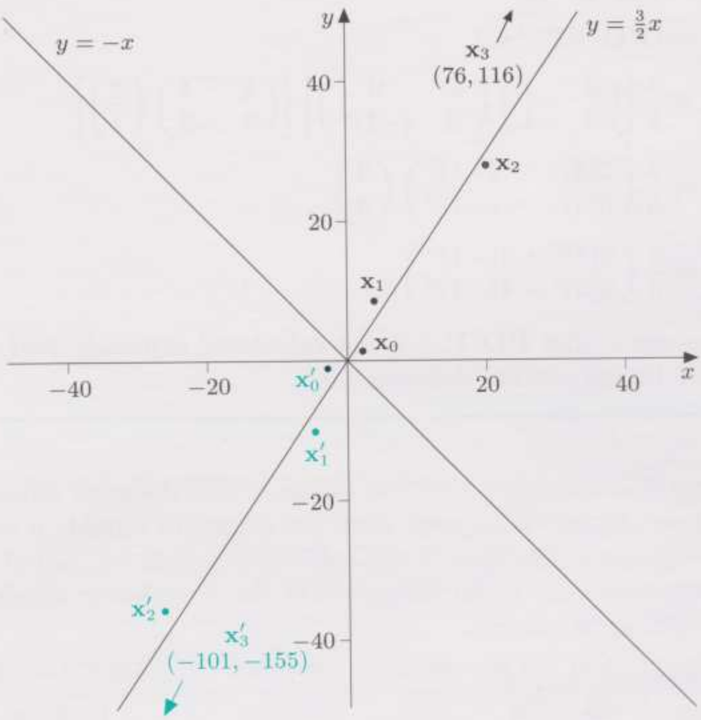


Figure 4.4 Two iteration sequences

In Figure 4.4, it appears that both iteration sequences are tending in the direction of the eigenline $y = \frac{3}{2}x$. (Note that 'tending in the direction of the eigenline $y = \frac{3}{2}x$' does not imply that the sequence is 'tending to the eigenline'. By property (b) of generalised scalings, the eigenvalue -1 associated with the eigenline $y = -x$ implies that each term of the sequence is the same distance from the eigenline $y = \frac{3}{2}x$, *and* consecutive terms lie on opposite sides of that line.)

We now show this conjecture to be true for the iteration sequence $\mathbf{x}_n$. In Example 4.1, we found a formula for the vector $\mathbf{x}_n$, namely

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 6(4)^n + 4(-1)^n \\ 9(4)^n - 4(-1)^n \end{pmatrix}.$$

Using this formula, we obtain the ratio

$$\frac{y_n}{x_n} = \frac{9(4)^n - 4(-1)^n}{6(4)^n + 4(-1)^n}.$$

The value of $-4(-1)^n$ is much smaller than $9(4)^n$ and the value of $4(-1)^n$ is much smaller than $6(4)^n$, so in a sense the terms involving 4^n 'dominate' for large values of n. To be precise, we divide both numerator and denominator by 4^n, to obtain

$$\frac{y_n}{x_n} = \frac{9 - 4(-\frac{1}{4})^n}{6 + 4(-\frac{1}{4})^n}.$$

As n gets large, $(-\frac{1}{4})^n$ tends to zero, so

$$\frac{y_n}{x_n} \to \frac{9-4(0)}{6+4(0)} = \frac{3}{2} \text{ as } n \to \infty.$$

The sequence $(-\frac{1}{4})^n$ is a geometric sequence with common ratio $-\frac{1}{4}$.

Since the ratio y_n/x_n tends to 1.5 (the gradient of the eigenline $y = \frac{3}{2}x$), the iteration sequence with initial point $(2,1)$ tends in the direction of this eigenline, as conjectured.

In fact, *any* iteration sequence generated by the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$$

which starts at a point not on an eigenline tends in the direction of the eigenline $y = \frac{3}{2}x$. However, the algebra required to prove this result is quite tedious, so instead a geometric explanation of this long-term behaviour is given.

Recall, from Subsection 3.3, that a linear transformation represented by a matrix which has two distinct non-zero eigenvalues is a generalised scaling. In this particular case, each time the linear transformation represented by the matrix $\mathbf{A}$ is applied, the points of the plane are scaled by the factor 4 in the direction of the line $y = \frac{3}{2}x$ and scaled by the factor -1 in the direction of the line $y = -x$. So successive points of the iteration sequence are scaled by the factor 4 in the direction of the line $y = \frac{3}{2}x$, and by the factor -1 in the direction of the line $y = -x$.

Using property (a) of generalised scalings, we deduce that successive points of an iteration sequence

See page 36.

- ◇ stay on the same side of the line $y = -x$, since $4 > 0$;
- ◇ move from one side of the line $y = \frac{3}{2}x$ to the other, since $-1 < 0$.

Figure 4.4 shows the two iteration sequences $\mathbf{x}_n$ and $\mathbf{x}'_n$ exhibiting this behaviour.

The following result enables the long-term behaviour of an iteration sequence generated by a matrix representing a generalised scaling to be described. Part (a) is a restatement of property (a) of generalised scalings for the terms of an iteration sequence. Part (b) contains criteria for deciding whether an iteration sequence moves away from $(0,0)$, as in the example discussed above, or whether it moves towards $(0,0)$. Part (c) describes the long-term behaviour of the ratio y_n/x_n in certain cases (the behaviour of this ratio was calculated in the example above).

In this result, the notation $\max\{a,b\}$ is used to denote the maximum of two real numbers a and b. For example,

$$\max\{2,-5\} = 2 \quad \text{and} \quad \max\{1,1\} = 1.$$

Iteration properties of generalised scalings

Let the linear transformation f be represented by a 2×2 matrix $\mathbf{A}$ that has two distinct non-zero eigenvalues k_1 and k_2 with corresponding eigenlines ℓ_1 and ℓ_2. Let (x_0, y_0) be a point of $\mathbb{R}^2$ which is not on an eigenline of $\mathbf{A}$, and let (x_n, y_n) be an iteration sequence generated by $\mathbf{A}$, with initial point (x_0, y_0).

(a) (i) If $k_1 > 0$, then all the points of (x_n, y_n) lie on the same side of ℓ_2 as (x_0, y_0).

(ii) If $k_1 < 0$, then the points of (x_n, y_n) alternate between opposite sides of ℓ_2.

(b) (i) If $\max\{|k_1|, |k_2|\} > 1$, then the sequence moves away from $(0, 0)$.

(ii) If $\max\{|k_1|, |k_2|\} < 1$, then the sequence moves towards $(0, 0)$.

(c) If $|k_1| > |k_2|$, then

$$\frac{y_n}{x_n} \to m \text{ as } n \to \infty,$$

where m is the gradient of ℓ_1.

If the eigenline ℓ_1 is $x = 0$, then

$$\frac{y_n}{x_n} \to \infty \text{ as } n \to \infty.$$

Remarks

1. Properties corresponding to (a) and (c) are obtained if k_1 and ℓ_2 are replaced by k_2 and ℓ_1, and vice versa.
2. If $|k_1| > |k_2|$, then k_1 is called the **dominant eigenvalue** and ℓ_1 is called the **dominant eigenline**. Property (c) is called the **Dominant Eigenvalue Property**.
3. The long-term behaviour specified in property (c) is often expressed as 'the sequence (x_n, y_n) tends in the direction of the line ℓ_1'.
4. Note that the case $\max\{|k_1|, |k_2|\} = 1$ is not covered by property (b), and the case $|k_1| = |k_2|$ is not covered by property (c).

In this section you will consider only matrices for which no eigenvalue has modulus less than 1. The next activity concerns such a case. (In Section 5, you will apply the properties to other cases.)

Activity 4.4 An iteration sequence

Consider the linear transformation represented by the matrix

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix}.$$

This matrix has eigenvalues -3 and 2, with corresponding eigenlines $y = -x$ and $y = 4x$. (You are *not* asked to establish these facts.)

Let $(2, 1)$ be the initial point of an iteration sequence (x_n, y_n) generated by $\mathbf{A}$. Note that $(2, 1)$ does not lie on either of the eigenlines of $\mathbf{A}$.

Use iteration properties (a), (b) and (c) of generalised scalings to describe the long-term behaviour of the sequence (x_n, y_n).

A solution is given on page 57.

An application to population models

The rest of this section will not be assessed.

This section concludes with a brief description of an application of matrix iteration to the long-term structure of subpopulations in population models. For example, a particular model of the UK population involves just two subpopulations – juveniles (those less than 15) and adults (the rest). Let J_n denote the number of juveniles and A_n denote the number of adults at the beginning of year n.

This population model, with n being the number of years after June 1990, was considered in MST121 Chapter B2, Section 3.

The relationship between these subpopulations in successive years is modelled by the recurrence relation

$$\begin{pmatrix} J_{n+1} \\ A_{n+1} \end{pmatrix} = \begin{pmatrix} 0.9326J_n + 0.0172A_n \\ 0.0666J_n + 0.9864A_n \end{pmatrix} \quad (n = 0, 1, 2, \ldots).$$

This equation can be expressed in matrix form as

$$\mathbf{p}_{n+1} = \mathbf{M}\mathbf{p}_n,$$

where $\mathbf{p}_n = \begin{pmatrix} J_n \\ A_n \end{pmatrix}$ and $\mathbf{M} = \begin{pmatrix} 0.9326 & 0.0172 \\ 0.0666 & 0.9864 \end{pmatrix}$. Thus $\mathbf{p}_n = \mathbf{M}^n\mathbf{p}_0$, where $\mathbf{p}_0$ represents (J_0, A_0).

It turns out that the matrix $\mathbf{M}$ has distinct eigenvalues:

$k_1 = 1.0027$, with eigenline $y = 4.0775x$;

$k_2 = 0.916\,27$, with eigenline $y = -0.949\,62x$.

(The values given are correct to five significant figures.)

Since $k_1 > k_2 > 0$, the dominant eigenvalue is k_1. According to the Dominant Eigenvalue Property, we have (for most initial subpopulations)

$$\frac{A_n}{J_n} \to 4.0775 \text{ as } n \to \infty.$$

This limit agrees with the result of numerical calculation given in MST121 Chapter B2.

So, according to the model, in the long-term there will be approximately four times as many adults as juveniles in the UK population.

This example is just one illustration of the many applications of matrices to the study of population models.

Now would be a good time to watch the *optional* band of Video Tape.

Now watch the video Band B(iv), 'Weaving spirals'.

Summary of Section 4

In this section we investigated iteration sequences for linear transformations represented by 2×2 matrices having two distinct eigenvalues.

This section has introduced:

◇ iteration sequences generated by a 2×2 matrix $\mathbf{A}$

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n \quad (n = 0, 1, 2, \ldots)$$

with initial point (x_0, y_0) represented by $\mathbf{x}_0$;

◇ the use of diagonalisation of $\mathbf{A}$ ($\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{A}$ has two distinct eigenvalues) to express the closed form $\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0$ of an iteration sequence in terms of n;

◇ results describing the long-term behaviour of iteration sequences generated by matrices representing generalised scalings.

Exercises for Section 4

Exercise 4.1

Determine the matrix $\mathbf{A}^n$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}$$

has diagonalisation

$$\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}.$$

Exercise 4.2

Consider the matrix $\mathbf{A} = \begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$, whose eigenvalues and eigenlines you calculated in Exercise 2.1(a).

(a) Which of the points $(1, 3)$, $(2, 2)$ and $(1, -5)$ lie on an eigenline of $\mathbf{A}$?

(b) Describe the long-term behaviour of iteration sequences (x_n, y_n) generated by $\mathbf{A}$ with each of the following initial points.

(i) $(1, 3)$

(ii) $(2, 2)$

(iii) $(1, -5)$

5 Iterating linear transformations with the computer

In this section you will need computer access and Computer Book B.

In Section 4, we looked at the long-term behaviour of iteration sequences for linear transformations which have two distinct non-zero eigenvalues. In this section, we consider such long-term behaviour for more general linear transformations. We also look at the iteration of some affine transformations. Recall from Chapter B2 that an affine transformation is a transformation of the form

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \mathbf{x} &\longmapsto \mathbf{A}\mathbf{x} + \mathbf{a}, \end{aligned}$$

where $\mathbf{A}$ is a 2×2 matrix and $\mathbf{a}$ is a vector.

We use the computer to investigate the iteration of linear and affine transformations, and discover how iteration can lead to pictures of natural objects such as ferns.

Refer to Computer Book B for the work in this section.

Summary of Section 5

This section has used the computer to help discover patterns occurring in the iteration of linear and affine transformations.

Summary of Chapter B3

In this chapter, you met eigenvalues, eigenlines and eigenvectors of a 2×2 matrix.

You saw that a matrix $\mathbf{A}$ which has two distinct eigenvalues can be diagonalised. The diagonal form enables powers of $\mathbf{A}$ to be calculated easily.

The long-term behaviour of iteration sequences generated by certain linear transformations was described.

Learning outcomes

You have been working towards the following learning outcomes.

Terms to know and use

Fixed point, invariant line, eigenvalue, eigenvector, eigenline, characteristic equation, eigenvector equation, diagonalising a matrix, generalised scaling, dominant eigenvalue, dominant eigenline.

Notation to know and use

$\max\{a, b\}$

Mathematical skills

- ◇ Determine the fixed points and invariant lines (through the origin) of certain linear transformations.
- ◇ Determine the eigenvalues of a 2×2 matrix using the characteristic equation.
- ◇ Find the eigenline corresponding to a given eigenvalue, and choose 'nice' eigenvectors.
- ◇ Write down the eigenvalues of diagonal and triangular matrices.
- ◇ Use the strategy for diagonalising a 2×2 matrix $\mathbf{A}$ which has two distinct eigenvalues, and hence write the matrix $\mathbf{A}$ in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.
- ◇ Compute powers of a 2×2 matrix which is expressible in the form $\mathbf{PDP}^{-1}$, where $\mathbf{D}$ is a diagonal matrix.
- ◇ Given a 2×2 matrix $\mathbf{A}$ with two distinct eigenvalues and an initial point (x_0, y_0), describe the long-term behaviour of the iteration sequence generated by $\mathbf{A}$.

Mathematical awareness

- ◇ Know that 2×2 matrices might have no eigenlines, one eigenline, two eigenlines or many eigenlines.

Mathcad skills

- ◇ Investigate iteration sequences of linear and affine transformations.

Summary of Block B

This block has introduced several key topics:

- ◇ iteration sequences, with recurrence relations of the form

  $$x_{n+1} = f(x_n) \quad (n = 0, 1, 2, \ldots),$$

  where f is a real function and x_0 is a given initial term;
- ◇ the Binomial Theorem;
- ◇ linear transformations $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ whose rules are of the form $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a 2×2 matrix;
- ◇ affine transformations $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ whose rules are of the form $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{a}$, where $\mathbf{A}$ is a 2×2 matrix and $\mathbf{a}$ is a vector;
- ◇ fixed points and invariant lines for linear transformations;
- ◇ eigenvalues, eigenlines and eigenvectors for 2×2 matrices;
- ◇ iteration sequences of points in $\mathbb{R}^2$, with recurrence relations of the form

  $$\mathbf{x}_{n+1} = f(\mathbf{x}_n) \quad (n = 0, 1, 2, \ldots),$$

  where f is a linear transformation or an affine transformation, and $\mathbf{x}_0$ is a given initial point.

In studying the long-term behaviour of iteration sequences generated by real functions, you saw the importance of the fixed points and p-cycles of the function f being iterated, and how these are classified using the gradient of the graph of f. The gradient of the graph of a real function is the fundamental concept in *calculus*, and you will learn much about this topic in Block C.

In studying linear transformations, you saw that the geometric nature of a linear transformation f can be related to properties of the matrix $\mathbf{A}$ that represents f. This relationship enabled us, for example, to describe the long-term behaviour of iteration sequences of points generated by certain linear transformations known as generalised scalings. You will see more about this relationship between linear transformations and their matrices in Block D.

Solutions to Activities

Solution 1.1

(a) The rotation $r_{3\pi/2}$ has only one fixed point, the origin.

(b) The reflection $q_{\pi/2}$ has a line of fixed points, the y-axis.

(c) The scaling with factors 1 and -1 has a line of fixed points, the x-axis.

(d) The y-shear with factor 4 has a line of fixed points, the y-axis.

Solution 1.2

(a) The rotation $r_{3\pi/2}$ has no invariant lines through the origin.

(b) The reflection $q_{\pi/2}$ has two invariant lines through the origin, the y-axis (which is a line of fixed points) and the x-axis.

(c) The scaling with factors 1 and -1 has two invariant lines through the origin, the x-axis and the y-axis.

(d) The y-shear with factor 4 has only one invariant line through the origin, the y-axis.

Solution 1.3

The matrix $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ represents the scaling with factors 1 and 3.

(a) The image of the point $(2, 0)$ under this scaling is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

so the point $(2, 0)$ is a fixed point of this scaling.

(b) Let $(0, c)$ be an arbitrary point on the y-axis. The image of the point $(0, c)$ under this scaling is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 3c \end{pmatrix}.$$

The point $(0, 3c)$ lies on the y-axis. Since $(0, c)$ is an arbitrary point on the y-axis, we have shown that every point on the y-axis has its image on that axis. Also, as c varies, this image ranges over the whole y-axis. Thus the y-axis is an invariant line of this scaling.

Solution 2.1

(a) The image under f of the arbitrary point $(c, -c)$ on the line $y = -x$ is given by

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} c \\ -c \end{pmatrix} = \begin{pmatrix} -c \\ c \end{pmatrix}.$$

The point $(-c, c)$ is also on the line $y = -x$, and, as c varies, ranges over the whole of that line. Thus this is an invariant line of f.

(b) The images are:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = -1\begin{pmatrix} 3 \\ -3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix} = -1\begin{pmatrix} -4 \\ 4 \end{pmatrix}.$$

(c) The image under f of the arbitrary point $(c, \frac{3}{2}c)$ on the line $y = \frac{3}{2}x$ is given by

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} c \\ \frac{3}{2}c \end{pmatrix} = \begin{pmatrix} 4c \\ 6c \end{pmatrix}.$$

The point $(4c, 6c)$ is also on the line $y = \frac{3}{2}x$, and, as c varies, ranges over the whole of that line. Thus this is an invariant line of f.

(d) The images are:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix} = 4\begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} = 4\begin{pmatrix} 1 \\ \frac{3}{2} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -2 \\ -3 \end{pmatrix} = \begin{pmatrix} -8 \\ -12 \end{pmatrix} = 4\begin{pmatrix} -2 \\ -3 \end{pmatrix}.$$

Solution 2.2

Only in part (c) is a full derivation of the characteristic equation given.

(a) The characteristic equation is

$$k^2 + k - 2 = 0.$$

This factorises as $(k + 2)(k - 1) = 0$, so the eigenvalues are -2 and 1.

(b) The characteristic equation is

$$k^2 - 6k + 23 = 0.$$

Using the quadratic formula, we obtain

$$k = \tfrac{1}{2}(6 \pm \sqrt{36 - 92}).$$

Since $36 - 92$ is negative, there are no real solutions to the characteristic equation; so this matrix has no eigenvalues.

(c) The characteristic equation is

$$k^2 - (-0.5 + 0.6)k + (-0.3) - (0.26) = 0;$$

that is,

$$k^2 - 0.1k - 0.56 = 0.$$

Using the quadratic formula, we obtain

$$\begin{aligned} k &= \tfrac{1}{2}(0.1 \pm \sqrt{2.25}) \\ &= \frac{0.1 \pm 1.5}{2}, \end{aligned}$$

so the eigenvalues are 0.8 and -0.7.

Solution 2.3

(a) In Activity 2.2, we saw that the eigenvalues for this matrix are -2 and 1, so we use these values for k in the eigenvector equation $\mathbf{A}\mathbf{x} = k\mathbf{x}$.

The eigenvector equation with eigenvalue -2,

$$\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -2 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$\begin{aligned} -x + 2y &= -2x, \\ x &= -2y. \end{aligned}$$

These equations both reduce to the equation $x + 2y = 0$, so we find that the eigenline corresponding to the eigenvalue -2 is $y = -\frac{1}{2}x$.

Any vector of the form $\begin{pmatrix} c \\ -\frac{1}{2}c \end{pmatrix}$, $c \neq 0$, is an eigenvector for this eigenline, so two such eigenvectors are $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} -4 \\ 2 \end{pmatrix}$.

The eigenvector equation with eigenvalue 1,

$$\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$\begin{aligned} -x + 2y &= x, \\ x &= y. \end{aligned}$$

These equations both reduce to the equation $y = x$, so the eigenline corresponding to the eigenvalue 1 is $y = x$.

Any vector of the form $\begin{pmatrix} c \\ c \end{pmatrix}$, $c \neq 0$, is an eigenvector for this eigenline, so two such eigenvectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ -2 \end{pmatrix}$.

(b) In Activity 2.2, we saw that the eigenvalues for this matrix are 0.8 and -0.7, so we use these values for k in the eigenvector equation $\mathbf{A}\mathbf{x} = k\mathbf{x}$.

The eigenvector equation with eigenvalue 0.8,

$$\begin{pmatrix} -0.5 & 1.3 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.8 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$\begin{aligned} -0.5x + 1.3y &= 0.8x, \\ 0.2x + 0.6y &= 0.8y. \end{aligned}$$

These equations both reduce to the equation $y = x$, so the eigenline corresponding to the eigenvalue 0.8 is $y = x$.

Any vector of the form $\begin{pmatrix} c \\ c \end{pmatrix}$, $c \neq 0$, is an eigenvector for this eigenline, so two such eigenvectors are $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$.

The eigenvector equation with eigenvalue -0.7,

$$\begin{pmatrix} -0.5 & 1.3 \\ 0.2 & 0.6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -0.7 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$\begin{aligned} -0.5x + 1.3y &= -0.7x, \\ 0.2x + 0.6y &= -0.7y. \end{aligned}$$

These equations both reduce to the equation $2x + 13y = 0$, so the eigenline corresponding to the eigenvalue -0.7 is $y = -\frac{2}{13}x$.

Any vector of the form $\begin{pmatrix} c \\ -\frac{2}{13}c \end{pmatrix}$, $c \neq 0$, is an eigenvector for this eigenline, so two such eigenvectors are $\begin{pmatrix} 13 \\ -2 \end{pmatrix}$ and $\begin{pmatrix} -26 \\ 4 \end{pmatrix}$.

Solution 2.4

The characteristic equation for this matrix is

$$k^2 - 3k - 4 = 0.$$

This factorises as $(k - 4)(k + 1) = 0$, so the eigenvalues are 4 and -1.

The eigenvector equation with eigenvalue 4,

$$\begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 4 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$\begin{aligned} -2x + 3y &= 4x, \\ -2x + 5y &= 4y. \end{aligned}$$

These equations both reduce to $-2x + y = 0$, so the eigenline corresponding to the eigenvalue 4 is $y = 2x$. An eigenvector for this eigenline is any vector of the form $\begin{pmatrix} c \\ 2c \end{pmatrix}$, $c \neq 0$, for example, the vector $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

The eigenvector equation with eigenvalue -1,

$$\begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -1 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-2x + 3y = -x,$$
$$-2x + 5y = -y.$$

These equations both reduce to $-x + 3y = 0$, so the eigenline corresponding to the eigenvalue -1 is $y = \frac{1}{3}x$. An eigenvector for this eigenline is any vector of the form $\begin{pmatrix} c \\ \frac{1}{3}c \end{pmatrix}$, $c \neq 0$, for example, the vector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Solution 2.5

(a) The characteristic equation is

$$k^2 + k - 6 = 0.$$

This factorises as $(k+3)(k-2) = 0$, so the eigenvalues are -3 and 2. These are the diagonal elements of the matrix.

(b) The characteristic equation for the matrix $\mathbf{D}$ is

$$k^2 - (a+d)k + ad = 0.$$

This factorises as $(k-a)(k-d) = 0$, so the eigenvalues are a and d, which are the diagonal elements of $\mathbf{D}$.

Solution 2.6

The eigenvector equation with eigenvalue -2,

$$\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -2\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-2x = -2x,$$
$$2y = -2y.$$

The first equation holds for every value of x and the second equation holds only for $y = 0$. So the eigenline corresponding to the eigenvalue -2 is $y = 0$; that is, it is the x-axis (as expected).

The eigenvector equation with eigenvalue 2,

$$\begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-2x = 2x,$$
$$2y = 2y.$$

The second equation holds for every value of y and the first equation holds only for $x = 0$. So the eigenline corresponding to the eigenvalue 2 is $x = 0$; that is, it is the y-axis (as expected).

Solution 2.7

(a) The matrix $\begin{pmatrix} 6 & 2 \\ 9 & 3 \end{pmatrix}$ has characteristic equation

$$k^2 - 9k = 0,$$

which factorises as $k(k-9) = 0$. Thus the matrix $\mathbf{A}$ has eigenvalues 0 and 9.

(b) The eigenvector equation with eigenvalue 0,

$$\begin{pmatrix} 6 & 2 \\ 9 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$6x + 2y = 0,$$
$$9x + 3y = 0.$$

These equations both reduce to $3x + y = 0$, which is the eigenline corresponding to the eigenvalue 0. This implies that every point on the line $3x + y = 0$ is mapped to the origin.

The eigenvector equation with eigenvalue 9,

$$\begin{pmatrix} 6 & 2 \\ 9 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 9\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$6x + 2y = 9x,$$
$$9x + 3y = 9y.$$

These equations both reduce to $3x - 2y = 0$, so this is the eigenline corresponding to the eigenvalue 9.

Solution 2.8

The characteristic equation for this matrix is

$$k^2 + 1 = 0.$$

This equation has no real solutions for k, so this matrix has no eigenvalues.

Solution 2.9

(a) The matrix $\mathbf{A}$ is triangular, so its eigenvalues are the elements on the leading diagonal. Thus this matrix has only one eigenvalue, $k = 1$. The eigenvector equation with eigenvalue 1,

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations $x = x$ and $x + y = y$. The first equation holds for every value of x (and y), but the second equation reduces to $x = 0$. Thus the y-axis is the only eigenline for this matrix.

(b) We saw in Section 1 that a y-shear has only one invariant line through the origin, the y-axis, so this result agrees with our geometric understanding of shears.

Solution 3.1

Step 1. The eigenvalues of $\mathbf{A}$ are solutions of the characteristic equation

$$k^2 - k - 6 = 0.$$

This equation factorises as $(k-3)(k+2) = 0$, so the eigenvalues are $k = 3$ and $k = -2$.

Step 2. Let $\mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$.

Step 3. The eigenvector equation with eigenvalue 3,

$$\begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations $x + 6y = 3x$ and $x = 3y$. These both reduce to $x - 3y = 0$. Thus $y = \frac{1}{3}x$ is the eigenline for the eigenvalue 3.

The eigenvector equation with eigenvalue -2,

$$\begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -2\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations $x + 6y = -2x$ and $x = -2y$. These both reduce to $x + 2y = 0$. Thus $y = -\frac{1}{2}x$ is the eigenline for the eigenvalue -2.

Step 4. For the eigenvalue 3, with eigenline $y = \frac{1}{3}x$, we can choose the eigenvector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

For the eigenvalue -2, with eigenline $y = -\frac{1}{2}x$, a good choice is $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Step 5. Hence we let $\mathbf{P} = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}$.

Step 6. Thus

$$\mathbf{P}^{-1} = \frac{1}{-5}\begin{pmatrix} -1 & -2 \\ -1 & 3 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}.$$

(We should now have

$$\begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}\frac{1}{5}\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix},$$

and this is easily checked, by multiplying out the right-hand side.)

Solution 3.2

The required matrix powers are:

$$\mathbf{A}^2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 6 \\ 9 & 10 \end{pmatrix};$$

$$\mathbf{A}^3 = \mathbf{A}\mathbf{A}^2 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 7 & 6 \\ 9 & 10 \end{pmatrix} = \begin{pmatrix} 25 & 26 \\ 39 & 38 \end{pmatrix};$$

$$\mathbf{D}^2 = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 16 & 0 \\ 0 & 1 \end{pmatrix};$$

$$\mathbf{D}^3 = \mathbf{D}\mathbf{D}^2 = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 16 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 64 & 0 \\ 0 & -1 \end{pmatrix}.$$

Solution 3.3

(a) $$\mathbf{P}^{-1} = \frac{1}{-5}\begin{pmatrix} 4 & -1 \\ -1 & -1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -4 & 1 \\ 1 & 1 \end{pmatrix}$$

(b) $$\begin{aligned}\mathbf{PDP}^{-1} &= \begin{pmatrix} -1 & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}\frac{1}{5}\begin{pmatrix} -4 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -3 & 8 \end{pmatrix}\begin{pmatrix} -4 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} -10 & 5 \\ 20 & 5 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix} = \mathbf{A}.\end{aligned}$$

(c) First we calculate $\mathbf{D}^4$:

$$\mathbf{D}^4 = \begin{pmatrix} (-3)^4 & 0 \\ 0 & 2^4 \end{pmatrix} = \begin{pmatrix} 81 & 0 \\ 0 & 16 \end{pmatrix}.$$

Thus

$$\begin{aligned}\mathbf{A}^4 &= \mathbf{PD}^4\mathbf{P}^{-1} \\ &= \begin{pmatrix} -1 & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 81 & 0 \\ 0 & 16 \end{pmatrix}\frac{1}{5}\begin{pmatrix} -4 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} -81 & 16 \\ 81 & 64 \end{pmatrix}\begin{pmatrix} -4 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 340 & -65 \\ -260 & 145 \end{pmatrix} \\ &= \begin{pmatrix} 68 & -13 \\ -52 & 29 \end{pmatrix}.\end{aligned}$$

Solution 3.4

(a)

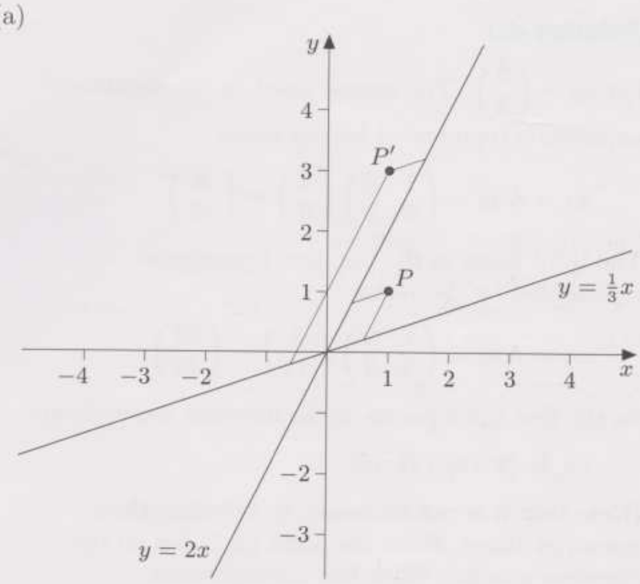


Figure S.1

(b) The image P' is given by

$$\begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

The construction in Figure S.1 agrees with this.

Solution 3.5

The matrix $\mathbf{A} = \begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix}$ has eigenvalues 4 and -1, with corresponding eigenlines $y = 2x$ and $y = \frac{1}{3}x$.

(a) (i) By property (a): since $4 > 0$, P' lies on the same side of the line $y = \frac{1}{3}x$ as P; since $-1 < 0$, P' lies on the opposite side of the line $y = 2x$ to P.

By property (b): since $|4| = 4$, the distance from P' to the line $y = \frac{1}{3}x$ is 4 times the distance from P to that line; since $|-1| = 1$, the distance from P' to the line $y = 2x$ is equal to the distance from P to that line.

(You should check these descriptions against Figure S.1.)

(ii) The required descriptions are the same as in part (i), but with P replaced by P', and P' replaced by P''. In particular, P'' is on the same side of the line $y = \frac{1}{3}x$ as P', but on the opposite side of the line $y = 2x$ to P'.

(b) The image P'' is given by

$$\begin{pmatrix} -2 & 3 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 13 \end{pmatrix}.$$

Note that $(7, 13)$ is indeed on the opposite side of the line $y = 2x$ to P'.

Solution 4.1

Let $\mathbf{x}_0 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. The second point in the iteration sequence is represented by the vector

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix}.$$

The third point in the iteration sequence is represented by the vector

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 8 \\ 12 \end{pmatrix} = \begin{pmatrix} 32 \\ 48 \end{pmatrix}.$$

So the first three points of the iteration sequence are

$$(2, 3), (8, 12), (32, 48).$$

(Note that it is not necessary to calculate these points as above. Since the point $(2, 3)$ lies on the eigenline $y = \frac{3}{2}x$, which has corresponding eigenvalue 4, the coordinates of each point in this iteration sequence are 4 times the coordinates of the previous point.)

Solution 4.2

(a) The initial point $(2, 3)$ lies on the eigenline $y = \frac{3}{2}x$, with eigenvalue 4. Hence, from the table, the iteration sequence moves away from $(0, 0)$, remaining on the same half of the line $y = \frac{3}{2}x$ as $(2, 3)$.

(b) The initial point $(2, -2)$ lies on the eigenline $y = -x$, with eigenvalue -1. Hence, from the table, the iteration sequence alternates between $(2, -2)$ and $(-2, 2)$.

Solution 4.3

(a) The second, third and fourth points in the iteration sequence with initial point $(-3, -1)$ are represented by the vectors:

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -3 \\ -1 \end{pmatrix} = \begin{pmatrix} -5 \\ -11 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -5 \\ -11 \end{pmatrix} = \begin{pmatrix} -27 \\ -37 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} -27 \\ -37 \end{pmatrix} = \begin{pmatrix} -101 \\ -155 \end{pmatrix}.$$

The first four points of the iteration sequence are $(-3, -1)$, $(-5, -11)$, $(-27, -37)$ and $(-101, -155)$.

(b) From Example 4.1, the matrix $\mathbf{A}^n$ is

$$\frac{1}{5}\begin{pmatrix} 2(4)^n + 3(-1)^n & 2(4)^n - 2(-1)^n \\ 3(4)^n - 3(-1)^n & 3(4)^n + 2(-1)^n \end{pmatrix}.$$

Now we calculate $\mathbf{x}'_n = \mathbf{A}^n\mathbf{x}'_0$:

$$\begin{aligned}
&\mathbf{x}'_n \\
&= \frac{1}{5}\begin{pmatrix} 2(4)^n + 3(-1)^n & 2(4)^n - 2(-1)^n \\ 3(4)^n - 3(-1)^n & 3(4)^n + 2(-1)^n \end{pmatrix}\begin{pmatrix} -3 \\ -1 \end{pmatrix} \\
&= \frac{1}{5}\begin{pmatrix} -6(4)^n - 9(-1)^n - 2(4)^n + 2(-1)^n \\ -9(4)^n + 9(-1)^n - 3(4)^n - 2(-1)^n \end{pmatrix} \\
&= \frac{1}{5}\begin{pmatrix} -8(4)^n - 7(-1)^n \\ -12(4)^n + 7(-1)^n \end{pmatrix}.
\end{aligned}$$

To find the 12^{th} point in this iteration sequence, we substitute $n = 11$ into the expression for $\mathbf{x}'_n$:

$$\begin{aligned}
\mathbf{x}'_{11} &= \frac{1}{5}\begin{pmatrix} -8(4)^{11} - 7(-1)^{11} \\ -12(4)^{11} + 7(-1)^{11} \end{pmatrix} \\
&= \frac{1}{5}\begin{pmatrix} -33\,554\,425 \\ -50\,331\,655 \end{pmatrix} \\
&= \begin{pmatrix} -6\,710\,885 \\ -10\,066\,331 \end{pmatrix}.
\end{aligned}$$

(c) The ratios are as follows (to 6 d.p.):

Point (x'_n, y'_n)	Ratio y'_n/x'_n
$(-3, -1)$	0.333 333
$(-5, -11)$	2.200 000
$(-27, -37)$	1.370 370
$(-101, -155)$	1.534 653
...	...
$(-6\,710\,885, -10\,066\,331)$	1.500 001

It appears that the ratio is tending to 1.5 once again.

Solution 4.4

By iteration property (a): since $2 > 0$, the points of the sequence (x_n, y_n) lie on the same side of the eigenline $y = -x$ as the initial point; since $-3 < 0$, the points alternate between opposite sides of the eigenline $y = 4x$.

By iteration property (b): since $\max\{|2|, |-3|\} = 3 > 1$, the sequence moves away from $(0, 0)$.

Since $|-3| > |2|$, the dominant eigenvalue is -3 and the dominant eigenline is $y = -x$. Hence, by iteration property (c) (the Dominant Eigenvalue Property),

$$\frac{y_n}{x_n} \to -1 \text{ as } n \to \infty.$$

Thus the long-term behaviour of the sequence may be described as follows. The sequence tends in the direction of the line $y = -x$, moving away from the origin. It stays on the same side of that line as the initial point, the points alternating between opposite sides of the line $y = 4x$.

The following figure, which you were *not* expected to produce, illustrates the first few terms of the sequence.

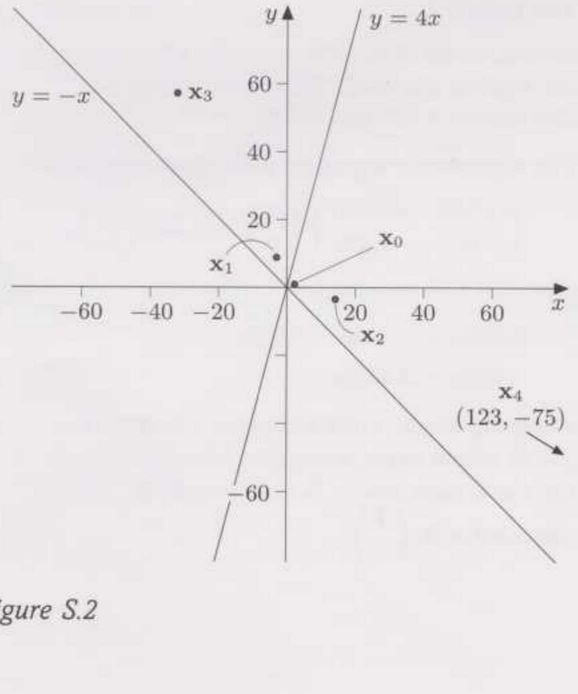


Figure S.2

Solutions to Exercises

Solution 1.1

(a) The rotation r_π has only one fixed point, the origin. Every line through the origin is an invariant line.

(b) The reflection $q_{\pi/4}$ has a line of fixed points, the line $y = x$. It has two invariant lines through the origin, the line $y = x$ (which is a line of fixed points) and the line $y = -x$.

(c) The scaling with factors 2 and -3 has only one fixed point, the origin. It has two invariant lines through the origin, the x-axis and the y-axis.

(d) The x-shear with factor 1 has a line of fixed points, the x-axis. It has only one invariant line through the origin, the x-axis (which is a line of fixed points).

Solution 1.2

(a) The matrix which represents the x-shear with factor 1 is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. To show that every point on the x-axis is a fixed point, consider an arbitrary point on the x-axis, say $(c, 0)$. The image of the point $(c, 0)$ is given by

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} c \\ 0 \end{pmatrix} = \begin{pmatrix} c \\ 0 \end{pmatrix},$$

so the point $(c, 0)$ is a fixed point. Since $(c, 0)$ is an arbitrary point on the x-axis, we have shown that every point on this axis is a fixed point.

(b) To show that the line $y = x$ is not an invariant line, we show that the image of a point on the line $y = x$ is not on the line $y = x$. Consider the point $(1, 1)$. The image of the point $(1, 1)$ is given by

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

The point $(2, 1)$ does not lie on the line $y = x$, so the line $y = x$ is not an invariant line of this shear. (In fact, no point on the line $y = x$ other than the origin has its image on the line $y = x$.)

Solution 2.1

(a) The matrix $\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$ has characteristic equation

$$k^2 - 8k + 7 = 0.$$

This factorises as $(k - 7)(k - 1) = 0$, so the eigenvalues of this matrix are 7 and 1.

The eigenvector equation with eigenvalue 7,

$$\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 7\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$6x + y = 7x,$$
$$5x + 2y = 7y.$$

These equations both reduce to the equation $y = x$, which is the eigenline corresponding to the eigenvalue 7. A corresponding eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

The eigenvector equation with eigenvalue 1,

$$\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$6x + y = x,$$
$$5x + 2y = y.$$

These equations both reduce to the equation $y = -5x$, which is the eigenline corresponding to the eigenvalue 1. A corresponding eigenvector is $\begin{pmatrix} 1 \\ -5 \end{pmatrix}$.

(b) The matrix $\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix}$ is a triangular matrix, so its eigenvalues are the elements on the leading diagonal. Thus this matrix has eigenvalues 3.535 and 3.396.

The eigenvector equation with eigenvalue 3.535,

$$\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 3.535\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$3.535x - 0.990y = 3.535x,$$
$$3.396y = 3.535y.$$

These equations both reduce to the equation $y = 0$, which is the eigenline corresponding to the eigenvalue 3.535. A corresponding eigenvector is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The eigenvector equation with eigenvalue 3.396,

$$\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3.396 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$3.535x - 0.990y = 3.396x,$$
$$3.396y = 3.396y.$$

The second equation holds for all values of y. The first equation reduces to

$$0.139x - 0.990y = 0; \quad \text{that is,} \quad y = \tfrac{139}{990}x,$$

which is the eigenline corresponding to the eigenvalue 3.396. A corresponding eigenvector is $\begin{pmatrix} 990 \\ 139 \end{pmatrix}$.

(c) The matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ has characteristic equation

$$k^2 - 5k = 0.$$

This factorises as $k(k-5) = 0$, so the eigenvalues of this matrix are 5 and 0. (This matrix represents a flattening.)

The eigenvector equation with eigenvalue 5,

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$x + 2y = 5x,$$
$$2x + 4y = 5y.$$

These equations both reduce to the equation $y = 2x$, which is the eigenline corresponding to the eigenvalue 5. A corresponding eigenvector is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

The eigenvector equation with eigenvalue 0,

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$x + 2y = 0,$$
$$2x + 4y = 0.$$

These equations both reduce to the equation $y = -\frac{1}{2}x$, which is the eigenline corresponding to the eigenvalue 0. (Thus all the points on this eigenline are mapped to the origin.) A corresponding eigenvector is $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Solution 2.2

(a) The matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ represents a scaling with factors 2 and 3. It should have two eigenlines, the x-axis and the y-axis, with corresponding eigenvalues 2 and 3, respectively.

The matrix $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ is a diagonal matrix, so its eigenvalues are 2 and 3.

The eigenvector equation with eigenvalue 2,

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$2x = 2x,$$
$$3y = 2y.$$

The first equation holds for all values of x, but the second equation reduces to $y = 0$. Thus the eigenline corresponding to the eigenvalue 2 is the line $y = 0$, which is the x-axis, as predicted.

The eigenvector equation with eigenvalue 3,

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$2x = 3x,$$
$$3y = 3y.$$

The second equation holds for all values of y, but the first equation reduces to $x = 0$. Thus the eigenline corresponding to the eigenvalue 3 is the line $x = 0$, which is the y-axis, as predicted.

(b) The matrix $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ represents an x-shear with factor 3. It should have one eigenline, the x-axis with corresponding eigenvalue 1 (since the points on the x-axis are fixed points).

The matrix $\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$ is a triangular matrix, so its eigenvalues are on the leading diagonal. Thus this matrix has only one eigenvalue, $k = 1$.

The eigenvector equation with eigenvalue 1,

$$\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$x + 3y = x,$$
$$y = y.$$

The second equation holds for all values of y, but the first equation reduces to $y = 0$. Thus the eigenline corresponding to the eigenvalue 1 is the line $y = 0$, which is the x-axis, as predicted.

(c) The matrix $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ represents a reflection in the line $y = -x$, which makes an angle of $3\pi/4$ with the positive x-axis. It should have two eigenlines. One is the axis of reflection $y = -x$ with corresponding eigenvalue 1, since this line consists of fixed points. The other is the line perpendicular to the axis of reflection, $y = x$ with corresponding eigenvalue -1, since points on this line are mapped to the same line on the opposite side of the origin but remain at the same distance from the origin.

The matrix $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$ has characteristic equation

$$k^2 - 1 = 0.$$

This factorises as $(k-1)(k+1) = 0$, giving the two eigenvalues 1 and -1.

The eigenvector equation with eigenvalue 1,

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-y = x,$$
$$-x = y.$$

These equations both reduce to $y = -x$, which is the eigenline corresponding to the eigenvalue 1, as predicted.

The eigenvector equation with eigenvalue -1,

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -1\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-y = -x,$$
$$-x = -y.$$

These equations both reduce to $y = x$, which is the eigenline corresponding to the eigenvalue -1, as predicted.

(d) The matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ corresponds to a rotation through an angle of π radians about the origin. It also corresponds to a uniform scaling with factor -1. It should have every line through the origin as an eigenline, with corresponding eigenvalue -1.

The matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is diagonal, so its eigenvalues are on the leading diagonal. Thus it has only one eigenvalue, $k = -1$.

The eigenvector equation with eigenvalue -1,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -1\begin{pmatrix} x \\ y \end{pmatrix},$$

gives the two equations

$$-x = -x,$$
$$-y = -y.$$

Both these equations hold for all values of x and y, so every line through the origin is an eigenline, as predicted.

Solution 3.1

(a) We follow the strategy, using the results from Exercise 2.1.

(i) Step 1. The eigenvalues of this matrix are 7 and 1.

Step 2. Let $\mathbf{D} = \begin{pmatrix} 7 & 0 \\ 0 & 1 \end{pmatrix}$.

Step 3. The eigenlines of this matrix are $y = x$ and $y = -5x$.

Step 4. An eigenvector for the eigenvalue 7 is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. An eigenvector for the eigenvalue 1 is $\begin{pmatrix} 1 \\ -5 \end{pmatrix}$.

Step 5. Hence we let $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & -5 \end{pmatrix}$.

Step 6. Thus

$$\mathbf{P}^{-1} = \frac{1}{-6}\begin{pmatrix} -5 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}.$$

Finally, we have

$$\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -5 \end{pmatrix}\begin{pmatrix} 7 & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}.$$

(ii) Step 1. The eigenvalues of this matrix are 3.535 and 3.396.

Step 2. Let $\mathbf{D} = \begin{pmatrix} 3.535 & 0 \\ 0 & 3.396 \end{pmatrix}$.

Step 3. The eigenlines of this matrix are $y = 0$ and $y = \frac{139}{990}x$.

Step 4. An eigenvector for the eigenvalue 3.535 is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. An eigenvector for the eigenvalue 3.396 is $\begin{pmatrix} 990 \\ 139 \end{pmatrix}$.

Step 5. Hence we let $\mathbf{P} = \begin{pmatrix} 1 & 990 \\ 0 & 139 \end{pmatrix}$.

Step 6. Thus

$$\mathbf{P}^{-1} = \frac{1}{139}\begin{pmatrix} 139 & -990 \\ 0 & 1 \end{pmatrix}.$$

Finally, we have

$$\begin{pmatrix} 3.535 & -0.990 \\ 0 & 3.396 \end{pmatrix} = \begin{pmatrix} 1 & 990 \\ 0 & 139 \end{pmatrix} \begin{pmatrix} 3.535 & 0 \\ 0 & 3.396 \end{pmatrix} \frac{1}{139} \begin{pmatrix} 139 & -990 \\ 0 & 1 \end{pmatrix}.$$

(iii) Step 1. The eigenvalues of this matrix are 5 and 0.

Step 2. Let $\mathbf{D} = \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix}$.

Step 3. The eigenlines of this matrix are $y = 2x$ and $y = -\frac{1}{2}x$.

Step 4. An eigenvector for the eigenvalue 5 is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. An eigenvector for the eigenvalue 0 is $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Step 5. Hence we let $\mathbf{P} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}$.

Step 6. Thus

$$\mathbf{P}^{-1} = \frac{1}{-5} \begin{pmatrix} -1 & -2 \\ -2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

Finally, we have

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}.$$

(b) We multiply the three matrices $\mathbf{PDP}^{-1}$ from part (a)(i):

$$\begin{aligned} \mathbf{PDP}^{-1} &= \begin{pmatrix} 1 & 1 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 7 & 1 \\ 7 & -5 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 36 & 6 \\ 30 & 12 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}, \end{aligned}$$

as required.

(c) First note that

$$\mathbf{D}^5 = \begin{pmatrix} 7^5 & 0 \\ 0 & 1^5 \end{pmatrix} = \begin{pmatrix} 16\,807 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} \mathbf{A}^5 &= \mathbf{PD}^5\mathbf{P}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 16\,807 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 16\,807 & 1 \\ 16\,807 & -5 \end{pmatrix} \begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 84\,036 & 16\,806 \\ 84\,030 & 16\,812 \end{pmatrix} \\ &= \begin{pmatrix} 14\,006 & 2801 \\ 14\,005 & 2802 \end{pmatrix}. \end{aligned}$$

Solution 3.2

The matrix $\mathbf{A} = \begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$ has eigenvalues 7 and 1, with corresponding eigenlines $y = x$ and $y = -5x$.

(a) By property (a): since $7 > 0$, P' lies on the same side of the line $y = -5x$ as P; since $1 > 0$, P' lies on the same side of the line $y = x$ as P.

By property (b): since $|7| = 7$, the distance from P' to the line $y = -5x$ is 7 times the distance from P to that line; since $|1| = 1$, the distance from P' to the line $y = x$ is equal to the distance from P to that line.

(b) The image P' is given by

$$\begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$

The locations of P and P' shown in Figure S.3 match the description in part (a).

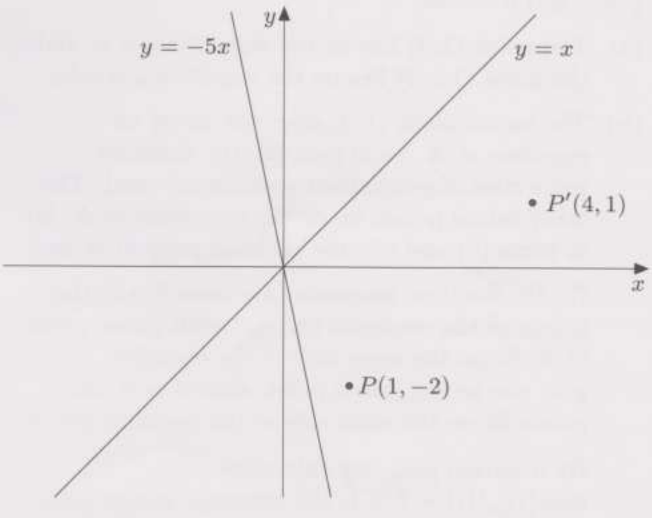


Figure S.3

As described in part (a): P and P' lie on the same side of the line $y = x$ and on the same side of the line $y = -5x$, they are equidistant from $y = x$, and P' is 7 times as distant from $y = -5x$ as P is.

Solution 4.1

Let $\mathbf{P} = \begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}$, $\mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$ and
$\mathbf{P}^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix}$.

Then $\mathbf{A} = \mathbf{PDP}^{-1}$, and

$$\begin{aligned}\mathbf{A}^n &= \mathbf{PD}^n\mathbf{P}^{-1} \\ &= \frac{1}{5}\begin{pmatrix} 3 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3^n & 0 \\ 0 & (-2)^n \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3(3^n) & 2(-2)^n \\ 3^n & -(-2)^n \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & -3 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 3(3^n)+2(-2)^n & 6(3^n)-6(-2)^n \\ 3^n-(-2)^n & 2(3^n)+3(-2)^n \end{pmatrix}.\end{aligned}$$

Solution 4.2

From Exercise 2.1(a), the matrix $\mathbf{A} = \begin{pmatrix} 6 & 1 \\ 5 & 2 \end{pmatrix}$ has eigenvalues 7 and 1 with corresponding eigenlines $y = x$ and $y = -5x$.

(a) The point $(2, 2)$ lies on the eigenline $y = x$, and the point $(1, -5)$ lies on the eigenline $y = -5x$.

(b) The initial point $(1, 3)$ does not lie on an eigenline of $\mathbf{A}$. So in part (i), the iteration properties of generalised scalings are used. The other initial points lie on the eigenlines of $\mathbf{A}$. So in parts (ii) and (iii) the table on page 41 is used.

(i) By iteration property (a): since $7 > 0$, the points of the sequence (x_n, y_n) with initial point $(1, 3)$ lie on the same side of the eigenline $y = -5x$ as the initial point; since $1 > 0$, the points lie on the same side of the eigenline $y = x$.

By iteration property (b): since $\max\{|7|, |1|\} = 7 > 1$, the sequence moves away from $(0, 0)$.

Since $|7| > |1|$, the dominant eigenvalue is 7 and the dominant eigenline is $y = x$. Hence by iteration property (c) (the Dominant Eigenvalue Property)

$$\frac{y_n}{x_n} \to 1 \text{ as } n \to \infty.$$

Thus the sequence tends in the direction of the line $y = x$, moving away from the origin. It stays on the same side of that line as the initial point and also on the same side of the line $y = -5x$.

(ii) The initial point $(2, 2)$ lies on the eigenline $y = x$, with eigenvalue 7. Hence, from the table, the iteration sequence moves away from $(0, 0)$, remaining on the same half of the line as $(2, 2)$.

(iii) The initial point $(1, -5)$ lies on the eigenline $y = -5x$, with eigenvalue 1. Hence, from the table, the iteration sequence is the constant sequence $(x_n, y_n) = (1, -5)$, for $n = 0, 1, 2, \ldots$.

Index

The Open University
ISBN 0 7492 4035 1